

(19) World Intellectual Property
Organization
International Bureau(43) International Publication Date
12 August 2004 (12.08.2004)

PCT

(10) International Publication Number
WO 2004/067707 A3(51) International Patent Classification⁷: C12N 15/00,
15/63, C12P 21/00, A01K 67/02770810 (US). FIORETTI, William, C. [US/US]; 2225 Lak-
eridge Road Drive, Grapevine, TX 76051 (US). CADD,
Gary, G. [US/US]; 501 Turner Road, Apartment 1111,
Grapevine, TX 76051 (US).(21) International Application Number:
PCT/US2003/041335(74) Agent: PRATT, John, S.; Kilpatrick Stockton LLP, Suite
2800, 1100 Peachtree Street, Atlanta, GA 30309 (US).(22) International Filing Date:
24 December 2003 (24.12.2003)(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR,
CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD,
GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,
MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU,
SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/441,377 21 January 2003 (21.01.2003) US
60/441,381 21 January 2003 (21.01.2003) US
60/441,392 21 January 2003 (21.01.2003) US
60/441,405 21 January 2003 (21.01.2003) US
60/441,447 21 January 2003 (21.01.2003) US
60/441,502 21 January 2003 (21.01.2003) US
10/609,019 26 June 2003 (26.06.2003) US(84) Designated States (regional): ARIPO patent (BW, GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE,
SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).(71) Applicants (for all designated States except US): TRANS-
GENRES, LLC [US/US]; 1750 Valley View Lane, Suite
100, Dallas, TX 75243 (US). THE BOARD OF SUPER-
VISORS OF LOUISIANA STATE UNIVERSITY AND
AGRICULTURAL AND MECHANICAL COLLEGE
[US/US]; P.O. Box 25055, Baton Rouge, LA 70894 (US).

Published:

— with international search report

(72) Inventors; and

(75) Inventors/Applicants (for US only): COOPER, Richard,
K. [US/US]; 111 Pecan Meadow Drive, Baton Rouge, LA(88) Date of publication of the international search report:
27 January 2005For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: ADMINISTRATION OF TRANSPOSON-BASED VECTORS TO REPRODUCTIVE ORGANS

(57) Abstract: Methods and compositions for the administration of transposon-based vectors to the reproductive organs of animals and the creation of transgenic animals. Preferred methods involve administration of the transposon-based vectors to the lumen of the oviduct of an avian, expression of a vector derived transgene in the avian, and deposition of the resultant polypeptide in an egg. This invention allows for large amounts of protein to be deposited in the egg.

WO 2004/067707 A3

5

1

10 **ADMINISTRATION OF TRANSPOSON-BASED VECTORS TO**
 REPRODUCTIVE ORGANS

 The U.S. Government has certain rights in this invention. The development of
this invention was partially funded by the United States Government under a HATCH
grant from the United States Department of Agriculture, partially funded by the
15 United States Government with Formula 1433 funds from the United States
Department of Agriculture and partially funded by the United States Government
under contract DAAD 19-02016 awarded by the Army.

FIELD OF THE INVENTION

20 The present invention relates generally to administration of a transposon-based
vector to the reproductive tract in an animal. The reproductive tract includes an
ovary, ova within an ovary, and an oviduct. Such administration results in
incorporation of a gene of interest contained in the vector in the ovary, the oviduct or
an ovum of the animal. In some embodiments, the present invention further includes
25 production of a protein encoded by the gene in an egg produced by the animal.

BACKGROUND OF THE INVENTION

 Transgenic animals are desirable for a variety of reasons, including their
potential as biological factories to produce desired molecules for pharmaceutical,
30 diagnostic and industrial uses. This potential is attractive to the industry due to the
inadequate capacity in facilities used for recombinant production of desired molecules
and the increasing demand by the pharmaceutical industry for use of these facilities.
Numerous attempts to produce transgenic animals have met several problems,
including low rates of gene incorporation and unstable gene incorporation.
35 Accordingly, improved gene technologies are needed for the development of
transgenic animals for the production of desired molecules.

Improved gene delivery technologies are also needed for the treatment of disease in animals and humans. Many diseases and conditions can be treated with gene-delivery technologies, which provide a gene of interest to a patient suffering from the disease or the condition. An example of such disease is Type 1 diabetes.

5 Type 1 diabetes is an autoimmune disease that ultimately results in destruction of the insulin producing β -cells in the pancreas. Although patients with Type 1 diabetes may be treated adequately with insulin injections or insulin pumps, these therapies are only partially effective. Insulin replacement, such as via insulin injection or pump administration, cannot fully reverse the defect in the vascular endothelium found in

10 the hyperglycemic state (Pieper et al., 1996. Diabetes Res. Clin. Pract. Suppl. S157-S162). In addition, hyper- and hypoglycemia occurs frequently despite intensive home blood glucose monitoring. Finally, careful dietary constraints are needed to maintain an adequate ratio of calories consumed. This often causes major psychosocial stress for many diabetic patients. Development of gene therapies

15 providing delivery of the insulin gene into the pancreas of diabetic patients could overcome many of these problems and result in improved life expectancy and quality of life.

Several of the prior art gene delivery technologies employed viruses that are associated with potentially undesirable side effects and safety concerns. The majority

20 of current gene-delivery technologies useful for gene therapy rely on virus-based delivery vectors, such as adeno and adeno-associated viruses, retroviruses, and other viruses, which have been attenuated to no longer replicate. (Kay, M.A., et al. 2001. Nature Medicine 7:33-40).

There are multiple problems associated with the use of viral vectors. Firstly,

25 they are not tissue-specific. In fact, a gene therapy trial using adenovirus was recently halted because the vector was present in the patient's sperm (Gene trial to proceed despite fears that therapy could change child's genetic makeup. The New York Times, December 23, 2001). Secondly, viral vectors are likely to be transiently incorporated, which necessitates re-treating a patient at specified time intervals. (Kay, M.A., et al. 2001. Nature Medicine 7:33-40). Thirdly, there is a concern that a viral-

30 based vector could revert to its virulent form and cause disease. Fourthly, viral-based vectors require a dividing cell for stable integration. Fifthly, viral-based vectors indiscriminately integrate into various cells, which can result in undesirable germline integration. Sixthly, the required high titers needed to achieve the desired effect have

35 resulted in the death of one patient and they are believed to be responsible for induction of cancer in a separate study. (Science, News of the Week, October 4, 2002).

Accordingly, what is needed is a new method to produce transgenic animals and humans with stably incorporated genes, in which the vector containing those genes does not cause disease or other unwanted side effects. There is also a need for DNA constructs that would be stably incorporated into the tissues and cells of animals and humans, including cells in the resting state that are not replicating. There is a further recognized need in the art for DNA constructs capable of delivering genes to specific tissues and cells of animals and humans.

When incorporating a gene of interest into an animal for the production of a desired protein or when incorporating a gene of interest in an animal or human for the treatment of a disease, it is often desirable to selectively activate incorporated genes using inducible promoters. These inducible promoters are regulated by substances either produced or recognized by the transcription control elements within the cell in which the gene is incorporated. In many instances, control of gene expression is desired in transgenic animals or humans so that incorporated genes are selectively activated at desired times and/or under the influence of specific substances. Accordingly, what is needed is a means to selectively activate genes introduced into the genome of cells of a transgenic animal or human. This can be taken a step further to cause incorporation to be tissue-specific, which prevents widespread gene incorporation throughout a patient's body (animal or human). This decreases the amount of DNA needed for a treatment, decreases the chance of incorporation in gametes, and targets gene delivery, incorporation, and expression to the desired tissue where the gene is needed to function. What is also needed is a rapid expression method for rapidly producing a protein or peptide of interest in eggs and milk of transgenic animals.

SUMMARY OF THE INVENTION

The present invention addresses the problems described above by providing new, effective and efficient compositions for producing transgenic animals and for treating disease in animals or humans. Transgenic animals include all egg-laying animals and milk-producing animals. Transgenic animals further include but are not limited to avians, fish, amphibians, reptiles, insects, mammals and humans. In another preferred embodiment, the animal is a milk-producing animal, including but not limited to bovine, porcine, ovine and equine animals. In a preferred embodiment, the animal is an avian animal. In another preferred embodiment, the animal is a mammal. Animals are made transgenic through administration of a composition comprising a transposon-based vector designed for incorporation of a gene of interest for production of a desired protein, together with an acceptable carrier. The

compositions of the present invention are introduced into the reproductive system of an animal. The compositions of the present invention are administered to a reproductive organ including, but not limited to, an oviduct, an ovary, or into the duct system of the mammary gland. The compositions of the present invention may be administered to a reproductive organ of an animal through the cloaca. The compositions of the present invention may be directly administered to a reproductive organ or can be administered to an artery leading to the reproductive organ. In a preferred embodiment, the compositions of the present invention are introduced into the the reproductive system of an avian animal. In another preferred embodiment, the compositions of the present invention are introduced into the the intramammary duct system of a mammal. A transfection reagent is optionally added to the composition before administration.

The transposon-based vectors of the present invention include a transposase, operably-linked to a first promoter, and a coding sequence for a protein or peptide of interest operably-linked to a second promoter, wherein the coding sequence for the protein or peptide of interest and its operably-linked promoter are flanked by transposase insertion sequences recognized by the transposase. The transposon-based vector also includes the following characteristics: a) one or more modified Kozak sequences at the 3' end of the first promoter to enhance expression of the transposase; b) modifications of the codons for the first several N-terminal amino acids of the transposase, wherein the nucleotide at the third base position of each codon is changed to an A or a T without changing the corresponding amino acid; c) addition of one or more stop codons to enhance the termination of transposase synthesis; and/or, d) addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene. In some embodiments, the effective polyA sequence is an avian optimized polyA sequence.

The present invention also provides for tissue-specific incorporation and/or expression of a gene of interest. Tissue-specific incorporation of a gene of interest may be achieved by placing the transposase gene under the control of a tissue-specific promoter, whereas tissue-specific expression of a gene of interest may be achieved by placing the gene of interest under the control of a tissue-specific promoter. In some embodiments, the gene of interest is transcribed under the influence of an ovalbumin, or other oviduct specific, promoter. Linking the gene of interest to an oviduct specific

promoter in an egg-laying animal results in synthesis of a desired molecule and deposition of the desired molecule in a developing egg.

The present invention advantageously produces a high number of transgenic animals having a gene of interest stably incorporated. In some embodiments wherein
5 the transposon-based vector is administered to the ovary, these transgenic animals successfully pass the desired gene to their progeny. Accordingly, the present invention can be used to obtain transgenic animals having the gene of interest incorporated into the germline through transfection of the ovary or the present invention can be used to obtain transgenic animals having the gene of interest
10 incorporated into the oviduct in a tissue-specific manner. Both types of transgenic animals of the present invention produce large amounts of a desired molecule encoded by the transgene. Transgenic egg-laying animals, particularly avians, produce large amounts of a desired protein that is deposited in the egg for rapid harvest and purification.

Any desired gene may be incorporated into the novel transposon-based vectors of the present invention in order to synthesize a desired molecule in the transgenic animals. Proteins, peptides and nucleic acids are preferred desired molecules to be
15 produced by the transgenic animals of the present invention. Particularly preferred proteins are antibody proteins and other immunopharmaceutical proteins.

This invention provides a composition useful for the production of transgenic
20 hens capable of producing substantially high amounts of a desired protein or peptide. Entire flocks of transgenic birds may be developed very quickly in order to produce industrial amounts of desired molecules. The present invention solves the problems inherent in the inadequate capacity of fermentation facilities used for bacterial
25 production of molecules and provides a more efficient and economical way to produce desired molecules. Accordingly, the present invention provides a means to produce large amounts of therapeutic, diagnostic and reagent molecules.

Transgenic chickens are excellent in terms of convenience and efficiency of
manufacturing molecules such as proteins and peptides. Starting with a single
30 transgenic rooster, thousands of transgenic offspring can be produced within a year. (In principle, up to forty million offspring could be produced in just three generations). Each transgenic female is expected to lay at least 250 eggs/year, each potentially containing hundreds of milligrams of the selected protein. Flocks of chickens numbering in the hundreds of thousands are readily handled through

established commercial systems. The technologies for obtaining eggs and fractionating them are also well known and widely accepted. Thus, for each therapeutic, diagnostic, or other protein of interest, large amounts of a substantially pure material can be produced at relatively low incremental cost.

5 A wide range of recombinant peptides and proteins can be produced in transgenic egg-laying animals. Enzymes, hormones, antibodies, growth factors, serum proteins, commodity proteins, biological response modifiers, peptides and designed proteins may all be made through practice of the present invention. For example, rough estimates suggest that it is possible to produce in bulk growth
10 hormone, insulin, or Factor VIII, and deposit them in egg whites, for an incremental cost in the order of one dollar per gram. At such prices it is feasible to consider administering such medical agents by inhalation or even orally, instead of through injection. Even if bioavailability rates through these avenues were low, the cost of a much higher effective-dose would not be prohibitive.

15 In one embodiment, the egg-laying transgenic animal is an avian. The method of the present invention may be used in avians including Ratites, Psittaciformes, Falconiformes, Piciformes, Strigiformes, Passeriformes, Coraciiformes, Ralliformes, Cuculiformes, Columbiformes, Galliformes, Anseriformes, and Herodiones. Preferably, the egg-laying transgenic animal is a poultry bird. More preferably, the
20 bird is a chicken, turkey, duck, goose or quail. Another preferred bird is a ratite, such as, an emu, an ostrich, a rhea, or a cassowary. Other preferred birds are partridge, pheasant, kiwi, parrot, parakeet, macaw, falcon, eagle, hawk, pigeon, cockatoo, song birds, jay bird, blackbird, finch, warbler, canary, toucan, mynah, or sparrow.

25 Accordingly, it is an object of the present invention to provide novel transposon-based vectors.

 It is another object of the present invention to provide novel transposon-based vectors that encode for the production of desired proteins or peptides in cells.

 It is an object of the present invention to produce transgenic animals through intraoviduct or intraovarian administration of a transposon-based vector.

30 Another object of the present invention is to produce transgenic animals through intraoviduct or intraovarian administration of a transposon-based vector, wherein the transgenic animals produce desired proteins or peptides.

It is further an object of the present invention to provide a method to produce transgenic animals through intraovarian administration of a transposon-based vector that are capable of producing transgenic progeny.

Yet another object of the present invention is to provide a method to produce
5 transgenic animals through intraoviduct or intraovarian administration of a transposon-based vector that are capable of producing a desired molecule, such as a protein, peptide or nucleic acid.

Another object of the present invention is to provide a method to produce transgenic animals through intraoviduct or intraovarian administration of a
10 transposon-based vector, wherein such administration results in modulation of endogenous gene expression.

It is yet another object of the present invention to provide a method to produce transgenic avians through intraoviduct or intraovarian administration of a transposon-based vector that are capable of producing proteins, peptides or nucleic acids.

It is another object of the present invention to produce transgenic animals
15 through intraoviduct or intraovarian administration of a transposon-based vector encoding an antibody or a fragment thereof.

Still another object of the present invention is to provide a method to produce transgenic avians through intraoviduct or intraovarian administration of a transposon-
20 based vector that are capable of producing proteins or peptides and depositing these proteins or peptides in the egg.

Another object of the present invention is to provide transgenic avians that contain a stably incorporated transgene.

Still another object of the present invention is to provide eggs containing
25 desired proteins or peptides encoded by a transgene incorporated into the transgenic avian that produces the egg.

It is further an object of the present invention to provide a method to produce transgenic milk-producing animals through administration of a transposon-based vector that are capable of producing proteins, peptides or nucleic acids.

Still another object of the present invention is to provide a method to produce transgenic milk-producing animals through administration of a transposon-based vector that are capable of producing proteins or peptides and depositing these proteins
30 or peptides in their milk.

Another object of the present invention is to provide transgenic milk-producing animals that contain a stably incorporated transgene.

Another object of the present invention is to provide transgenic milk-producing animals that are capable of producing proteins or peptides and depositing
5 these proteins or peptides in their milk.

Yet another object of the present invention is to provide milk containing desired molecules encoded by a transgene incorporated into the transgenic milk-producing animals that produce the milk.

Still another object of the present invention is to provide milk containing
10 desired proteins or peptides encoded by a transgene incorporated into the transgenic milk-producing animals that produce the milk.

An advantage of the present invention is that transgenic animals are produced with higher efficiencies than observed in the prior art.

Another advantage of the present invention is that these transgenic animals
15 possess high copy numbers of the transgene.

Another advantage of the present invention is that the transgenic animals produce large amounts of desired molecules encoded by the transgene.

Still another advantage of the present invention is that desired molecules are produced by the transgenic animals much more efficiently and economically than
20 prior art methods, thereby providing a means for large scale production of desired molecules, particularly proteins and peptides.

Yet another advantage of the present invention is that the desired proteins and peptides are produced rapidly after making animals transgenic through introduction of the vectors of the present invention.

25 These and other objects, features and advantages of the present invention will become apparent after a review of the following detailed description of the disclosed embodiments and claims.

BRIEF DESCRIPTION OF THE FIGURES

30 Figure 1 depicts schematically a transposon-based vector containing a transposase operably linked to a first promoter and a gene of interest operably-linked to a second promoter, wherein the gene of interest and its operably-linked promoter are flanked by insertion sequences (IS) recognized by the transposase. "Pro"

designates a promoter. In this and subsequent figures, the size of the actual nucleotide sequence is not necessarily proportionate to the box representing that sequence.

Figure 2 depicts schematically a transposon-based vector for targeting deposition of a polypeptide in an egg white wherein Ov pro is the ovalbumin promoter, Ov protein is the ovalbumin protein and PolyA is a polyadenylation sequence. The TAG sequence includes a spacer sequence, the gp41 hairpin loop from HIV I and a protease cleavage site.

Figure 3 depicts schematically a transposon-based vector for targeting deposition of a polypeptide in an egg white wherein Ovo pro is the ovomucoid promoter and Ovo SS is the ovomucoid signal sequence. The TAG sequence includes a spacer, the gp41 hairpin loop from HIV I and a protease cleavage site.

Figure 4 depicts schematically a transposon based-vector for expression of an RNAi molecule. "Tet_i pro" indicates a tetracycline inducible promoter whereas "pro" indicates the pro portion of a prepro sequence as described herein. "Ovgen" indicates approximately 60 base pairs of an ovalbumin gene, "Ovotrans" indicates approximately 60 base pairs of an ovotransferrin gene and "Ovomucin" indicates approximately 60 base pairs of an ovomucin gene.

Figure 5 is a picture of an SDS-PAGE gel wherein a pooled fraction of an isolated proinsulin fusion protein was run in lanes 4 and 6. Lanes 1 and 10 of the gel contain molecular weight standards, lanes 2 and 8 contain non-transgenic chicken egg white, and lanes 3, 5, 7 and 9 are blank.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a new, effective and efficient method of producing transgenic animals, particularly egg-laying animals and milk-producing animals, through administration of a composition comprising a transposon-based vector designed for incorporation of a gene of interest and production of a desired molecule. The transposon-based vectors are administered to a reproductive organ including, but not limited to, an oviduct, an ovary, or into the duct system of the mammary gland. The vectors may be directly administered to a reproductive organ or can be administered to an artery leading to the reproductive organ or to a lymph system proximate to the cells to be genetically altered. The vectors may be administered to a reproductive organ of an animal through the cloaca. One method of direct administration is by injection, and in one embodiment, the lumen of the

magnum of the oviduct is injected with a transposon-based vector. Another method of direct administration is by injection, and in one embodiment, the lumen of the infundibulum of the oviduct is injected with a transposon-based vector. A preferred intrarterial administration is an administration into an artery that supplies the oviduct or the ovary. In some embodiments, administration of the transposon-based vector to an oviduct or an artery that leads to the oviduct results in incorporation of the vector into the epithelial and/or secretory cells of the oviduct. In other embodiments, administration of the transposon-based vector to an ovary or an artery that leads to the ovary or a lymphatic system proximal to the ovary results in incorporation of the vector into an oocyte or a germinal disk inside the ovary.

Definitions

It is to be understood that as used in the specification and in the claims, "a" or "an" can mean one or more, depending upon the context in which it is used. Thus, for example, reference to "a cell" can mean that at least one cell can be utilized.

The term "antibody" is used interchangeably with the term "immunoglobulin" and is defined herein as a protein synthesized by an animal or a cell of the immune system in response to the presence of a foreign substance commonly referred to as an "antigen" or an "immunogen". The term antibody includes fragments of antibodies. Antibodies are characterized by specific affinity to a site on the antigen, wherein the site is referred to an "antigenic determinant" or an "epitope". Antigens can be naturally occurring or artificially engineered. Artificially engineered antigens include, but are not limited to, small molecules, such as small peptides, attached to haptens such as macromolecules, for example proteins, nucleic acids, or polysaccharides. Artificially designed or engineered variants of naturally occurring antibodies and artificially designed or engineered antibodies not occurring in nature are all included in the current definition. Such variants include conservatively substituted amino acids and other forms of substitution as described in the section concerning proteins and polypeptides.

As used herein, the term "egg-laying animal" includes all amniotes such as birds, turtles, lizards and monotremes. Monotremes are egg-laying mammals and include the platypus and echidna. The term "bird" or "fowl," as used herein, is defined as a member of the Aves class of animals which are characterized as warm-blooded, egg-laying vertebrates primarily adapted for flying. Avians include, without limitation, Ratites, Psittaciformes, Falconiformes, Piciformes, Strigiformes,

Passeriformes, Coraciformes, Ralliformes, Cuculiformes, Columbiformes, Galliformes, Anseriformes, and Herodiones. The term "Ratite," as used herein, is defined as a group of flightless, mostly large, running birds comprising several orders and including the emus, ostriches, kiwis, and cassowaries. The term "Psittaciformes",
5 as used herein, includes parrots and refers to a monofamilial order of birds that exhibit zygodactylism and have a strong hooked bill. A "parrot" is defined as any member of the avian family Psittacidae (the single family of the Psittaciformes), distinguished by the short, stout, strongly hooked beak. Avians include all poultry birds, especially chickens, geese, turkeys, ducks and quail. The term "chicken" as used herein denotes
10 chickens used for table egg production, such as egg-type chickens, chickens reared for public meat consumption, or broilers, and chickens reared for both egg and meat production ("dual-purpose" chickens). The term "chicken" also denotes chickens produced by primary breeder companies, or chickens that are the parents, grandparents, great-grandparents, etc. of those chickens reared for public table egg,
15 meat, or table egg and meat consumption.

The term "egg" is defined herein as including a large female sex cell enclosed in a porous, calcarous or leathery shell, produced by birds and reptiles. The term "ovum" is defined as a female gamete, and is also known as an egg. Therefore, egg
20 production in all animals other than birds and reptiles, as used herein, is defined as the production and discharge of an ovum from an ovary, or "ovulation". Accordingly, it is to be understood that the term "egg" as used herein is defined as a large female sex cell enclosed in a porous, calcarous or leathery shell, when a bird or reptile produces it, or it is an ovum when it is produced by all other animals.

The term "milk-producing animal" refers herein to mammals including, but
25 not limited to, bovine, ovine, porcine, equine, and primate animals. Milk-producing animals include but are not limited to cows, llamas, camels, goats, reindeer, zebu, water buffalo, yak, horses, pigs, rabbits, non-human primates, and humans.

The term "gene" is defined herein to include a coding region for a protein, peptide or polypeptide.

30 The term "transgenic animal" refers to an animal having at least a portion of the transposon-based vector DNA incorporated into its DNA. While a transgenic animal includes an animal wherein the transposon-based vector DNA is incorporated into the germline DNA, a transgenic animal also includes an animal having DNA in one or more cells that contain a portion of the transposon-based vector DNA for any

period of time. In a preferred embodiment, a portion of the transposon-based vector comprises a gene of interest. More preferably, the gene of interest is incorporated into the animal's DNA for a period of at least five days, more preferably the reproductive life of the animal, and most preferably the life of the animal. In a further preferred
5 embodiment, the animal is an avian.

The term "vector" is used interchangeably with the terms "construct", "DNA construct" and "genetic construct" to denote synthetic nucleotide sequences used for manipulation of genetic material, including but not limited to cloning, subcloning, sequencing, or introduction of exogenous genetic material into cells, tissues or
10 organisms, such as birds. It is understood by one skilled in the art that vectors may contain synthetic DNA sequences, naturally occurring DNA sequences, or both. The vectors of the present invention are transposon-based vectors as described herein.

When referring to two nucleotide sequences, one being a regulatory sequence, the term "operably-linked" is defined herein to mean that the two sequences are
15 associated in a manner that allows the regulatory sequence to affect expression of the other nucleotide sequence. It is not required that the operably-linked sequences be directly adjacent to one another with no intervening sequence(s).

The term "regulatory sequence" is defined herein as including promoters, enhancers and other expression control elements such as polyadenylation sequences,
20 matrix attachment sites, insulator regions for expression of multiple genes on a single construct, ribosome entry/attachment sites, introns that are able to enhance expression, and silencers.

Transposon-Based Vectors

While not wanting to be bound by the following statement, it is believed that
25 the nature of the DNA construct is an important factor in successfully producing transgenic animals. The "standard" types of plasmid and viral vectors that have previously been almost universally used for transgenic work in all species, especially avians, have low efficiencies and may constitute a major reason for the low rates of transformation previously observed. The DNA (or RNA) constructs previously used
30 often do not integrate into the host DNA, or integrate only at low frequencies. Other factors may have also played a part, such as poor entry of the vector into target cells. The present invention provides transposon-based vectors that can be administered to an animal that overcome the prior art problems relating to low transgene integration frequencies. Two preferred transposon-based vectors of the present invention in

which a transposase, gene of interest and other polynucleotide sequences may be introduced are termed pTnMCS (SEQ ID NO:2) and pTnMod (SEQ ID NO:3).

The transposon-based vectors of the present invention produce integration frequencies an order of magnitude greater than has been achieved with previous
5 vectors. More specifically, intratesticular injections performed with a prior art transposon-based vector (described in U.S. Patent No. 5,719,055) resulted in 41% sperm positive roosters whereas intratesticular injections performed with the novel transposon-based vectors of the present invention resulted in 77% sperm positive roosters. Actual frequencies of integration were estimated by either or both
10 comparative strength of the PCR signal from the sperm and histological evaluation of the testes and sperm by quantitative PCR.

The transposon-based vectors of the present invention include a transposase gene operably-linked to a first promoter, and a coding sequence for a desired protein or peptide operably-linked to a second promoter, wherein the coding sequence for the
15 desired protein or peptide and its operably-linked promoter are flanked by transposase insertion sequences recognized by the transposase. The transposon-based vector also includes one or more of the following characteristics: a) one or more modified Kozak sequences comprising ACCATG (SEQ ID NO:1) at the 3' end of the first promoter to enhance expression of the transposase; b) modifications of the codons for the first
20 several N-terminal amino acids of the transposase, wherein the third base of each codon was changed to an A or a T without changing the corresponding amino acid; c) addition of one or more stop codons to enhance the termination of transposase synthesis; and/or, d) addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene. The transposon-
25 based vector may additionally or alternatively include one or more of the following Kozak sequences at the 3' end of any promoter, including the promoter operably-linked to the transposase: ACCATGG (SEQ ID NO:4), AAGATGT (SEQ ID NO:5), ACGATGA (SEQ ID NO:6), AAGATGG (SEQ ID NO:7), GACATGA (SEQ ID NO:8), ACCATGA (SEQ ID NO:9), and ACCATGA (SEQ ID NO:10), ACCATGT
30 (SEQ ID NO:52).

Figure 1 shows a schematic representation of several components of the transposon-based vector. The present invention further includes vectors containing more than one gene of interest, wherein a second or subsequent gene of interest is operably-linked to the second promoter or to a different promoter. It is also to be

understood that the transposon-based vectors shown in the Figures are representative of the present invention and that the order of the vector elements may be different than that shown in the Figures, that the elements may be present in various orientations, and that the vectors may contain additional elements not shown in the

5 Figures.

Transposases and Insertion Sequences

In a further embodiment of the present invention, the transposase found in the transposase-based vector is an altered target site (ATS) transposase and the insertion sequences are those recognized by the ATS transposase. However, the transposase
10 located in the transposase-based vectors is not limited to a modified ATS transposase and can be derived from any transposase. Transposases known in the prior art include those found in AC7, Tn5SEQ1, Tn916, Tn951, Tn1721, Tn 2410, Tn1681, Tn1, Tn2, Tn3, Tn4, Tn5, Tn6, Tn9, Tn10, Tn30, Tn101, Tn903, Tn501, Tn1000 ($\gamma\delta$), Tn1681, Tn2901, AC transposons, Mp transposons, Spm transposons, En transposons, Dotted
15 transposons, Mu transposons, Ds transposons, dSpm transposons and I transposons. According to the present invention, these transposases and their regulatory sequences are modified for improved functioning as follows: a) the addition one or more modified Kozak sequences comprising ACCATG (SEQ ID NO:1) at the 3' end of the promoter operably-linked to the transposase; b) a change of the codons for the first
20 several amino acids of the transposase, wherein the third base of each codon was changed to an A or a T without changing the corresponding amino acid; c) the addition of one or more stop codons to enhance the termination of transposase synthesis; and/or, d) the addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene.

25 Although not wanting to be bound by the following statement, it is believed that the modifications of the first several N-terminal codons of the transposase gene increase transcription of the transposase gene, in part, by increasing strand dissociation. It is preferable that between approximately 1 and 20, more preferably 3 and 15, and most preferably between 4 and 12 of the first N-terminal codons of the
30 transposase are modified such that the third base of each codon is changed to an A or a T without changing the encoded amino acid. In one embodiment, the first ten N-terminal codons of the transposase gene are modified in this manner. It is also preferred that the transposase contain mutations that make it less specific for preferred

insertion sites and thus increases the rate of transgene insertion as discussed in U.S. Patent No. 5,719,055.

In some embodiments, the transposon-based vectors are optimized for expression in a particular host by changing the methylation patterns of the vector DNA. For example, prokaryotic methylation may be reduced by using a methylation deficient organism for production of the transposon-based vector. The transposon-based vectors may also be methylated to resemble eukaryotic DNA for expression in a eukaryotic host.

Transposases and insertion sequences from other analogous eukaryotic transposon-based vectors that can also be modified and used are, for example, the *Drosophila* P element derived vectors disclosed in U.S. Patent No. 6,291,243; the *Drosophila* mariner element described in Sherman et al. (1998); or the sleeping beauty transposon. See also Hackett et al. (1999); D. Lampe et al., 1999. Proc. Natl. Acad. Sci. USA, 96:11428-11433; S. Fischer et al., 2001. Proc. Natl. Acad. Sci. USA, 98:6759-6764; L. Zagoraiou et al., 2001. Proc. Natl. Acad. Sci. USA, 98:11474-11478; and D. Berg et al. (Eds.), *Mobile DNA*, Amer. Soc. Microbiol. (Washington, D.C., 1989). However, it should be noted that bacterial transposon-based elements are preferred, as there is less likelihood that a eukaryotic transposase in the recipient species will recognize prokaryotic insertion sequences bracketing the transgene.

Many transposases recognize different insertion sequences, and therefore, it is to be understood that a transposase-based vector will contain insertion sequences recognized by the particular transposase also found in the transposase-based vector. In a preferred embodiment of the invention, the insertion sequences have been shortened to about 70 base pairs in length as compared to those found in wild-type transposons that typically contain insertion sequences of well over 100 base pairs.

While the examples provided below incorporate a "cut and insert" Tn10 based vector that is destroyed following the insertion event, the present invention also encompasses the use of a "rolling replication" type transposon-based vector. Use of a rolling replication type transposon allows multiple copies of the transposon/transgene to be made from a single transgene construct and the copies inserted. This type of transposon-based system thereby provides for insertion of multiple copies of a transgene into a single genome. A rolling replication type transposon-based vector may be preferred when the promoter operably-linked to gene of interest is endogenous to the host cell and present in a high copy number or highly expressed. However, use

of a rolling replication system may require tight control to limit the insertion events to non-lethal levels. Tn1, Tn2, Tn3, Tn4, Tn5, Tn9, Tn21, Tn501, Tn551, Tn951, Tn1721, Tn2410 and Tn2603 are examples of a rolling replication type transposon, although Tn5 could be both a rolling replication and a cut and insert type transposon.

5 Stop Codons and PolyA Sequences

In one embodiment, the transposon-based vector contains two stop codons operably-linked to the transposase and/or to the gene of interest. In an alternate embodiment, one stop codon of UAA or UGA is operably linked to the transposase and/or to the gene of interest.

10 As used herein an "effective polyA sequence" refers to either a synthetic or non-synthetic sequence that contains multiple and sequential nucleotides containing an adenine base (an A polynucleotide string) and that increases expression of the gene to which it is operably-linked. A polyA sequence may be operably-linked to any gene in the transposon-based vector including, but not limited to, a transposase gene and a
15 gene of interest. A preferred polyA sequence is optimized for use in the host animal or human. In one embodiment, the polyA sequence is optimized for use in an avian species and more specifically, a chicken. An avian optimized polyA sequence generally contains a minimum of 40 base pairs, preferably between approximately 40 and several hundred base pairs, and more preferably approximately 75 base pairs that
20 precede the A polynucleotide string and thereby separate the stop codon from the A polynucleotide string. In one embodiment of the present invention, the polyA sequence comprises a conalbumin polyA sequence as provided in SEQ ID NO:11 and as taken from GenBank accession # Y00407, base pairs 10651-11058. In another
25 embodiment, the polyA sequence comprises a synthetic polynucleotide sequence shown in SEQ ID NO:12. In yet another embodiment, the polyA sequence comprises an avian optimized polyA sequence provided in SEQ ID NO:13. A chicken optimized polyA sequence may also have a reduced amount of CT repeats as compared to a synthetic polyA sequence.

30 It is a surprising discovery of the present invention that such an avian optimized poly A sequence increases expression of a polynucleotide to which it is operably-linked in an avian as compared to a non-avian optimized polyA sequence. Accordingly, the present invention includes methods of or increasing incorporation of a gene of interest wherein the gene of interest resides in a transposon-based vector containing a transposase gene and wherein the transposase gene is operably linked to

an avian optimized polyA sequence. The present invention also includes methods of increasing expression of a gene of interest in an avian that includes administering a gene of interest to the avian, wherein the gene of interest is operably-linked to an avian optimized polyA sequence. An avian optimized polyA nucleotide string is defined herein as a polynucleotide containing an A polynucleotide string and a minimum of 40 base pairs, preferably between approximately 40 and several hundred base pairs, and more preferably approximately 60 base pairs that precede the A polynucleotide string. The present invention further provides transposon-based vectors containing a gene of interest or transposase gene operably linked to an avian optimized polyA sequence.

Promoters and Enhancers

The first promoter operably-linked to the transposase gene and the second promoter operably-linked to the gene of interest can be a constitutive promoter or an inducible promoter. Constitutive promoters include, but are not limited to, immediate early cytomegalovirus (CMV) promoter, herpes simplex virus 1 (HSV1) immediate early promoter, SV40 promoter, lysozyme promoter, early and late CMV promoters, early and late HSV promoters, β -actin promoter, tubulin promoter, Rous-Sarcoma virus (RSV) promoter, and heat-shock protein (HSP) promoter. Inducible promoters include tissue-specific promoters, developmentally-regulated promoters and chemically inducible promoters. Examples of tissue-specific promoters include the glucose 6 phosphate (G6P) promoter, vitellogenin promoter, ovalbumin promoter, ovomucoid promoter, conalbumin promoter, ovotransferrin promoter, prolactin promoter, kidney uromodulin promoter, and placental lactogen promoter. In one embodiment, the vitellogenin promoter includes a polynucleotide sequence of SEQ ID NO:14. The G6P promoter sequence may be deduced from a rat G6P gene untranslated upstream region provided in GenBank accession number U57552.1. Examples of developmentally-regulated promoters include the homeobox promoters and several hormone induced promoters. Examples of chemically inducible promoters include reproductive hormone induced promoters and antibiotic inducible promoters such as the tetracycline inducible promoter and the zinc-inducible metallothionein promoter.

Other inducible promoter systems include the Lac operator repressor system inducible by IPTG (isopropyl beta-D-thiogalactoside) (Cronin, A. et al. 2001. Genes

and Development, v. 15), ecdysone-based inducible systems (Hoppe, U. C. et al. 2000. Mol. Ther. 1:159-164); estrogen-based inducible systems (Braselmann, S. et al. 1993. Proc. Natl. Acad. Sci. 90:1657-1661); progesterone-based inducible systems using a chimeric regulator, GLVP, which is a hybrid protein consisting of the GAL4
5 binding domain and the herpes simplex virus transcriptional activation domain, VP16, and a truncated form of the human progesterone receptor that retains the ability to bind ligand and can be turned on by RU486 (Wang, et al. 1994. Proc. Natl. Acad. Sci. 91:8180-8184); CID-based inducible systems using chemical inducers of dimerization (CIDs) to regulate gene expression, such as a system wherein rapamycin induces
10 dimerization of the cellular proteins FKBP12 and FRAP (Belshaw, P. J. et al. 1996. J. Chem. Biol. 3:731-738; Fan, L. et al. 1999. Hum. Gene Ther. 10:2273-2285; Shariat, S.F. et al. 2001. Cancer Res. 61:2562-2571; Spencer, D.M. 1996. Curr. Biol. 6:839-847). Chemical substances that activate the chemically inducible promoters can be administered to the animal containing the transgene of interest via any method known
15 to those of skill in the art.

Other examples of cell or tissue-specific and constitutive promoters include but are not limited to smooth-muscle SM22 promoter, including chimeric SM22alpha/telokin promoters (Hoggatt A.M. et al., 2002. Circ Res. 91(12):1151-9); ubiquitin C promoter (Biochim Biophys Acta, 2003. Jan. 3;1625(1):52-63); Hsf2
20 promoter; murine COMP (cartilage oligomeric matrix protein) promoter; early B cell-specific mb-1 promoter (Sigvardsson M., et al., 2002. Mol. Cell Biol. 22(24):8539-51); prostate specific antigen (PSA) promoter (Yoshimura I. et al., 2002, J. Urol. 168(6):2659-64); exorh promoter and pineal expression-promoting element (Asaoka Y., et al., 2002. Proc. Natl. Acad. Sci. 99(24):15456-61); neural and liver ceramidase
25 gene promoters (Okino N. et al., 2002. Biochem. Biophys. Res. Commun. 299(1):160-6); PSP94 gene promoter/enhancer (Gabril M.Y. et al., 2002. Gene Ther. 9(23):1589-99); promoter of the human FAT/CD36 gene (Kuriki C., et al., 2002. Biol. Pharm. Bull. 25(11):1476-8); VL30 promoter (Staplin W.R. et al., 2002. Blood
October 24, 2002); and, IL-10 promoter (Brenner S., et al., 2002. J. Biol. Chem.
30 December 18, 2002).

Examples of avian promoters include, but are not limited to, promoters controlling expression of egg white proteins, such as ovalbumin, ovotransferrin (conalbumin), ovomucoid, lysozyme, ovomucin, g2 ovoglobulin, g3 ovoglobulin, ovoflavoprotein, ovostatin (ovomacroglobin), cystatin, avidin, thiamine-binding

protein, glutamyl aminopeptidase minor glycoprotein 1, minor glycoprotein 2; and promoters controlling expression of egg-yolk proteins, such as vitellogenin, very low-density lipoproteins, low density lipoprotein, cobalamin-binding protein, riboflavin-binding protein, biotin-binding protein (Awade, 1996. Z. Lebensm. Unters. Forsch. 202:1-14). An advantage of using the vitellogenin promoter is that it is active during the egg-laying stage of an animal's life-cycle, which allows for the production of the protein of interest to be temporally connected to the import of the protein of interest into the egg yolk when the protein of interest is equipped with an appropriate targeting sequence. In some embodiments, the avian promoter is an oviduct-specific promoter. As used herein, the term "oviduct-specific promoter" includes, but is not limited to, ovalbumin; ovotransferrin (conalbumin); ovomucoid; 01, 02, 03, 04 or 05 avidin; ovomucin; g2 ovoglobulin; g3 ovoglobulin; ovoflavoprotein; and ovostatin (ovomacroglobin) promoters.

When germline transformation occurs via intraovarian administration, liver-specific promoters may be operably-linked to the gene of interest to achieve liver-specific expression of the transgene. Liver-specific promoters of the present invention include, but are not limited to, the following promoters, vitellogenin promoter, G6P promoter, cholesterol-7-alpha-hydroxylase (CYP7A) promoter, phenylalanine hydroxylase (PAH) promoter, protein C gene promoter, insulin-like growth factor I (IGF-I) promoter, bilirubin UDP-glucuronosyltransferase promoter, aldolase B promoter, furin promoter, metallothioneine promoter, albumin promoter, and insulin promoter.

Also included in the present invention are promoters that can be used to target expression of a protein of interest into the milk of a milk-producing animal including, but not limited to, β lactoglobulin promoter, whey acidic protein promoter, lactalbumin promoter and casein promoter.

When germline transformation occurs via intraovarian administration, immune system-specific promoters may be operably-linked to the gene of interest to achieve immune system-specific expression of the transgene. Accordingly, promoters associated with cells of the immune system may also be used. Acute phase promoters such as interleukin (IL)-1 and IL-2 may be employed. Promoters for heavy and light chain Ig may also be employed. The promoters of the T cell receptor components CD4 and CD8, B cell promoters and the promoters of CR2 (complement receptor type

2) may also be employed. Immune system promoters are preferably used when the desired protein is an antibody protein.

Also included in this invention are modified promoters/enhancers wherein elements of a single promoter are duplicated, modified, or otherwise changed. In one embodiment, steroid hormone-binding domains of the ovalbumin promoter are moved from about -6.5 kb to within approximately the first 1000 base pairs of the gene of interest. Modifying an existing promoter with promoter/enhancer elements not found naturally in the promoter, as well as building an entirely synthetic promoter, or drawing promoter/enhancer elements from various genes together on a non-natural backbone, are all encompassed by the current invention.

Accordingly, it is to be understood that the promoters contained within the transposon-based vectors of the present invention may be entire promoter sequences or fragments of promoter sequences. For example, in one embodiment, the promoter operably linked to a gene of interest is an approximately 900 base pair fragment of a chicken ovalbumin promoter (SEQ ID NO:15). The constitutive and inducible promoters contained within the transposon-based vectors may also be modified by the addition of one or more modified Kozak sequences of ACCATG (SEQ ID NO:1).

As indicated above, the present invention includes transposon-based vectors containing one or more enhancers. These enhancers may or may not be operably-linked to their native promoter and may be located at any distance from their operably-linked promoter. A promoter operably-linked to an enhancer and a promoter modified to eliminate repressive regulatory effects are referred to herein as an "enhanced promoter." The enhancers contained within the transposon-based vectors are preferably enhancers found in birds, and more preferably, an ovalbumin enhancer, but are not limited to these types of enhancers. In one embodiment, an approximately 675 base pair enhancer element of an ovalbumin promoter is cloned upstream of an ovalbumin promoter with 300 base pairs of spacer DNA separating the enhancer and promoter. In one embodiment, the enhancer used as a part of the present invention comprises base pairs 1-675 of a chicken ovalbumin enhancer from GenBank accession #S82527.1. The polynucleotide sequence of this enhancer is provided in SEQ ID NO:16.

Also included in some of the transposon-based vectors of the present invention are cap sites and fragments of cap sites. In one embodiment, approximately 50 base pairs of a 5' untranslated region wherein the capsite resides are added on the 3' end of

an enhanced promoter or promoter. An exemplary 5' untranslated region is provided in SEQ ID NO:17. A putative cap-site residing in this 5' untranslated region preferably comprises the polynucleotide sequence provided in SEQ ID NO:18.

In one embodiment of the present invention, the first promoter operably-linked to the transposase gene is a constitutive promoter and the second promoter operably-linked to the gene of interest is a tissue-specific promoter. In the second embodiment, use of the first constitutive promoter allows for constitutive activation of the transposase gene and incorporation of the gene of interest into virtually all cell types, including the germline of the recipient animal. Although the gene of interest is incorporated into the germline generally, the gene of interest may only be expressed in a tissue-specific manner. A transposon-based vector having a constitutive promoter operably-linked to the transposase gene can be administered by any route, and in one embodiment, the vector is administered to an ovary, to an artery leading to the ovary or to a lymphatic system or fluid proximal to the ovary.

It should be noted that cell- or tissue-specific expression as described herein does not require a complete absence of expression in cells or tissues other than the preferred cell or tissue. Instead, "cell-specific" or "tissue-specific" expression refers to a majority of the expression of a particular gene of interest in the preferred cell or tissue, respectively.

When incorporation of the gene of interest into the germline is not preferred, the first promoter operably-linked to the transposase gene can be a tissue-specific promoter. For example, transfection of a transposon-based vector containing a transposase gene operably-linked to an oviduct specific promoter such as the ovalbumin promoter provides for activation of the transposase gene and incorporation of the gene of interest in the cells of the oviduct but not into the germline and other cells generally. In this embodiment, the second promoter operably-linked to the gene of interest can be a constitutive promoter or an inducible promoter. In a preferred embodiment, both the first promoter and the second promoter are an ovalbumin promoter. In embodiments wherein tissue-specific expression or incorporation is desired, it is preferred that the transposon-based vector is administered directly to the tissue of interest, to an artery leading to the tissue of interest or to fluids surrounding the tissue of interest. In a preferred embodiment, the tissue of interest is the oviduct and administration is achieved by direct injection into the oviduct or an artery leading to the oviduct. In a further preferred embodiment, administration is achieved by

direct injection into the lumen of the magnum or the infundibulum of the oviduct. Indirect administration to the oviduct may occur through the cloaca.

Accordingly, cell specific promoters may be used to enhance transcription in selected tissues. In birds, for example, promoters that are found in cells of the fallopian tube, such as ovalbumin, conalbumin, ovomucoid and/or lysozyme, are used in the vectors to ensure transcription of the gene of interest in the epithelial cells and tubular gland cells of the fallopian tube, leading to synthesis of the desired protein encoded by the gene and deposition into the egg white. In mammals, promoters specific for the epithelial cells of the alveoli of the mammary gland, such as prolactin, insulin, beta lactoglobulin, whey acidic protein, lactalbumin, casein, and/or placental lactogen, are used in the design of vectors used for transfection of these cells for the production of desired proteins for deposition into the milk. In liver cells, the G6P promoter may be employed to drive transcription of the gene of interest for protein production. Proteins made in the liver of birds may be delivered to the egg yolk.

In order to achieve higher or more efficient expression of the transposase gene, the promoter and other regulatory sequences operably-linked to the transposase gene may be those derived from the host. These host specific regulatory sequences can be tissue specific as described above or can be of a constitutive nature. For example, an avian actin promoter and its associated polyA sequence can be operably-linked to a transposase in a transposase-based vector for transfection into an avian. Examples of other host specific promoters that could be operably-linked to the transposase include the myosin and DNA or RNA polymerase promoters.

Directing Sequences

In some embodiments of the present invention, the gene of interest is operably-linked to a directing sequence or a sequence that provides proper conformation to the desired protein encoded by the gene of interest. As used herein, the term "directing sequence" refers to both signal sequences and targeting sequences. An egg directing sequence includes, but is not limited to, an ovomucoid signal sequence, an ovalbumin signal sequence, a cecropin pre pro signal sequence, and a vitellogenin targeting sequence. The term "signal sequence" refers to an amino acid sequence, or the polynucleotide sequence that encodes the amino acid sequence, that directs the protein to which it is linked to the endoplasmic reticulum in a eukaryote, and more preferably the translocational pores in the endoplasmic reticulum, or the plasma membrane in a prokaryote, or mitochondria, such as for the purpose of gene

therapy for mitochondrial diseases. Signal and targeting sequences can be used to direct a desired protein into, for example, the milk, when the transposon-based vectors are administered to a milk-producing animal.

Signal sequences can also be used to direct a desired protein into, for example, a secretory pathway for incorporation into the egg yolk or the egg white, when the transposon-based vectors are administered to a bird or other egg-laying animal. One example of such a transposon-based vector is provided in Figure 3 wherein the gene of interest is operably linked to the ovomucoid signal sequence. The present invention also includes a gene of interest operably-linked to a second gene containing a signal sequence. An example of such an embodiment is shown in Figure 2 wherein the gene of interest is operably-linked to the ovalbumin gene that contains an ovalbumin signal sequence. Other signal sequences that can be included in the transposon-based vectors include, but are not limited to the ovotransferrin and lysozyme signal sequences. In one embodiment, the signal sequence is an ovalbumin signal sequence including a sequence shown in SEQ ID NO:19. In another embodiment, the signal sequence is a modified ovalbumin signal sequence including a sequence shown in SEQ ID NO:20 or SEQ ID NO:21.

As also used herein, the term "targeting sequence" refers to an amino acid sequence, or the polynucleotide sequence encoding the amino acid sequence, which amino acid sequence is recognized by a receptor located on the exterior of a cell. Binding of the receptor to the targeting sequence results in uptake of the protein or peptide operably-linked to the targeting sequence by the cell. One example of a targeting sequence is a vitellogenin targeting sequence that is recognized by a vitellogenin receptor (or the low density lipoprotein receptor) on the exterior of an oocyte. In one embodiment, the vitellogenin targeting sequence includes the polynucleotide sequence of SEQ ID NO:22. In another embodiment, the vitellogenin targeting sequence includes all or part of the vitellogenin gene. Other targeting sequences include VLDL and Apo E, which are also capable of binding the vitellogenin receptor. Since the ApoE protein is not endogenously expressed in birds, its presence may be used advantageously to identify birds carrying the transposon-based vectors of the present invention.

Genes of Interest Encoding Desired Proteins

A gene of interest selected for stable incorporation is designed to encode any desired protein or peptide or to regulate any cellular response. In some embodiments,

the desired proteins or peptides are deposited in an egg or in milk. It is to be understood that the present invention encompasses transposon-based vectors containing multiple genes of interest. The multiple genes of interest may each be operably-linked to a separate promoter and other regulatory sequence(s) or may all be operably-linked to the same promoter and other regulatory sequences(s). In one embodiment, multiple gene of interest are linked to a single promoter and other regulatory sequence(s) and each gene of interest is separated by a cleavage site or a pro portion of a signal sequence. A gene of interest may contain modifications of the codons for the first several N-terminal amino acids of the gene of interest, wherein the third base of each codon is changed to an A or a T without changing the corresponding amino acid.

Protein and peptide hormones are a preferred class of proteins in the present invention. Such protein and peptide hormones are synthesized throughout the endocrine system and include, but are not limited to, hypothalamic hormones and hypophysiotropic hormones, anterior, intermediate and posterior pituitary hormones, pancreatic islet hormones, hormones made in the gastrointestinal system, renal hormones, thymic hormones, parathyroid hormones, adrenal cortical and medullary hormones. Specifically, hormones that can be produced using the present invention include, but are not limited to, chorionic gonadotropin, corticotropin, erythropoietin, glucagons, IGF-1, oxytocin, platelet-derived growth factor, calcitonin, follicle-stimulating hormone, luteinizing hormone, thyroid-stimulating hormone, insulin, gonadotropin-releasing hormone and its analogs, vasopressin, octreotide, somatostatin, prolactin, adrenocorticotrophic hormone, antidiuretic hormone, thyrotropin-releasing hormone (TRH), growth hormone-releasing hormone (GHRH), dopamine, melatonin, thyroxine (T_4), parathyroid hormone (PTH), glucocorticoids such as cortisol, mineralocorticoids such as aldosterone, androgens such as testosterone, adrenaline (epinephrine), noradrenaline (norepinephrine), estrogens such as estradiol, progesterone, glucagons, calcitriol, calciferol, atrial-natriuretic peptide, gastrin, secretin, cholecystokinin (CCK), neuropeptide Y, ghrelin, PYY₃₋₃₆, angiotensinogen, thrombopoietin, and leptin. By using appropriate polynucleotide sequences, species-specific hormones may be made by transgenic animals.

In one embodiment of the present invention, the gene of interest is a proinsulin gene and the desired molecule is insulin. Proinsulin consists of three parts: a C-peptide and two strands of amino acids (the alpha and beta chains) that later become

linked together to form the insulin molecule. Figures 2 and 3 are schematics of transposon-based vector constructs containing a proinsulin gene operably-linked to an ovalbumin promoter and ovalbumin protein or an ovomucoid promoter and ovomucoid signal sequence, respectively. In these embodiments, proinsulin is expressed in the oviduct tubular gland cells and then deposited in the egg white. One example of a proinsulin polynucleotide sequence is shown in SEQ ID NO:23, wherein the C-peptide cleavage site spans from Arg at position 31 to Arg at position 65.

Serum proteins including lipoproteins such as high density lipoprotein (HDL), HDL-Milano and low density lipoprotein, albumin, clotting cascade factors, factor VIII, factor IX, fibrinogen, and globulins are also included in the group of desired proteins of the present invention. Immunoglobulins are one class of desired globulin molecules and include but are not limited to IgG, IgM, IgA, IgD, IgE, IgY, lambda chains, kappa chains and fragments thereof; Fc fragments, and Fab fragments. Desired antibodies include, but are not limited to, naturally occurring antibodies, human antibodies, humanized antibodies, and hybrid antibodies. Genes encoding modified versions of naturally occurring antibodies or fragments thereof and genes encoding artificially designed antibodies or fragments thereof may be incorporated into the transposon-based vectors of the present invention. Desired antibodies also include antibodies with the ability to bind specific ligands, for example, antibodies against proteins associated with cancer-related molecules, such as anti-her 2, or anti-CA125. Accordingly, the present invention encompasses a transposon-based vector containing one or more genes encoding a heavy immunoglobulin (Ig) chain and a light Ig chain. Further, more than one gene encoding for more than one antibody may be administered in one or more transposon-based vectors of the present invention. In this manner, an egg may contain more than one type of antibody in the egg white, the egg yolk or both. In one embodiment, a transposon-based vector contains a heavy Ig chain and a light Ig chain, both operably linked to a promoter.

Antibodies used as therapeutic reagents include but are not limited to antibodies for use in cancer immunotherapy against specific antigens, or for providing passive immunity to an animal or a human against an infectious disease or a toxic agent. Antibodies used as diagnostic reagents include, but are not limited to antibodies that may be labeled and detected with a detector, for example antibodies with a fluorescent label attached that may be detected following exposure to specific wavelengths. Such labeled antibodies may be primary antibodies directed to a

specific antigen, for example, rhodamine-labeled rabbit anti-growth hormone, or may be labeled secondary antibodies, such as fluorescein-labeled goat-anti chicken IgG. Such labeled antibodies are known to one of ordinary skill in the art. Labels useful for attachment to antibodies are also known to one of ordinary skill in the art. Some of these labels are described in the "Handbook of Fluorescent Probes and Research Products", ninth edition, Richard P. Haugland (ed) Molecular Probes, Inc. Eugene, OR), which is incorporated herein in its entirety.

Antibodies produced with using the present invention may be used as laboratory reagents for numerous applications including radioimmunoassay, western blots, dot blots, ELISA, immunoaffinity columns and other procedures requiring antibodies as known to one of ordinary skill in the art. Such antibodies include primary antibodies, secondary antibodies and tertiary antibodies, which may be labeled or unlabeled.

Antibodies that may be made with the practice of the present invention include, but are not limited to primary antibodies, secondary antibodies, designer antibodies, anti-protein antibodies, anti-peptide antibodies, anti-DNA antibodies, anti-RNA antibodies, anti-hormone antibodies, anti-hypophysiotropic peptides, antibodies against non-natural antigens, anti-anterior pituitary hormone antibodies, anti-posterior pituitary hormone antibodies, anti-venom antibodies, anti-tumor marker antibodies, antibodies directed against epitopes associated with infectious disease, including, anti-viral, anti-bacterial, anti-protozoal, anti-fungal, anti-parasitic, anti-receptor, anti-lipid, anti-phospholipid, anti-growth factor, anti-cytokine, anti-monokine, anti-idiotypic, and anti-accessory (presentation) protein antibodies. Antibodies made with the present invention, as well as light chains or heavy chains, may also be used to inhibit enzyme activity.

Antibodies that may be produced using the present invention include, but are not limited to, antibodies made against the following proteins: Bovine γ -Globulin, Serum; Bovine IgG, Plasma; Chicken γ -Globulin, Serum; Human γ -Globulin, Serum; Human IgA, Plasma; Human IgA₁, Myeloma; Human IgA₂, Myeloma; Human IgA₂, Plasma; Human IgD, Plasma; Human IgE, Myeloma; Human IgG, Plasma; Human IgG, Fab Fragment, Plasma; Human IgG, F(ab')₂ Fragment, Plasma; Human IgG, Fc Fragment, Plasma; Human IgG₁, Myeloma; Human IgG₂, Myeloma; Human IgG₃, Myeloma; Human IgG₄, Myeloma; Human IgM, Myeloma; Human IgM, Plasma; Human Immunoglobulin, Light Chain κ , Urine; Human Immunoglobulin, Light

Chains κ and λ , Plasma; Mouse γ -Globulin, Serum; Mouse IgG, Serum; Mouse IgM, Myeloma; Rabbit γ -Globulin, Serum; Rabbit IgG, Plasma; and Rat γ -Globulin, Serum. In one embodiment, the transposon-based vector comprises the coding sequence of light and heavy chains of a murine monoclonal antibody that shows
 5 specificity for human seminoprotein (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively).

A further non-limiting list of antibodies that recognize other antibodies is as follows: Anti-Chicken IgG, heavy (H) & light (L) Chain Specific (Sheep); Anti-Goat γ -Globulin (Donkey); Anti-Goat IgG, Fc Fragment Specific (Rabbit); Anti-Guinea Pig
 10 γ -Globulin (Goat); Anti-Human Ig, Light Chain, Type κ Specific; Anti-Human Ig, Light Chain, Type λ Specific; Anti-Human IgA, α -Chain Specific (Goat); Anti-Human IgA, Fab Fragment Specific; Anti-Human IgA, Fc Fragment Specific; Anti-Human IgA, Secretory; Anti-Human IgE, ϵ -Chain Specific (Goat); Anti-Human IgE, Fc Fragment Specific; Anti-Human IgG, Fc Fragment Specific (Goat); Anti-Human
 15 IgG, γ -Chain Specific (Goat); Anti-Human IgG, Fc Fragment Specific; Anti-Human IgG, Fd Fragment Specific; Anti-Human IgG, H & L Chain Specific (Goat); Anti-Human IgG₁, Fc Fragment Specific; Anti-Human IgG₂, Fc Fragment Specific; Anti-Human IgG₂, Fd Fragment Specific; Anti-Human IgG₃, Hinge Specific; Anti-Human IgG₄, Fc Fragment Specific; Anti-Human IgM, Fc Fragment Specific; Anti-Human
 20 IgM, μ -Chain Specific; Anti-Mouse IgE, ϵ -Chain Specific; Anti-Mouse γ -Globulin (Goat); Anti-Mouse IgG, γ -Chain Specific (Goat); Anti-Mouse IgG, γ -Chain Specific (Goat) F(ab')₂ Fragment; Anti-Mouse IgG, H & L Chain Specific (Goat); Anti-Mouse IgM, μ -Chain Specific (Goat); Anti-Mouse IgM, H & L Chain Specific (Goat); Anti-Rabbit γ -Globulin (Goat); Anti-Rabbit IgG, Fc Fragment Specific (Goat); Anti-Rabbit
 25 IgG, H & L Chain Specific (Goat); Anti-Rat γ -Globulin (Goat); Anti-Rat IgG, H & L Chain Specific; Anti-Rhesus Monkey γ -Globulin (Goat); and, Anti-Sheep IgG, H & L Chain Specific.

Another non-limiting list of the antibodies that may be produced using the present invention is provided in product catalogs of companies such as Phoenix
 30 Pharmaceuticals, Inc. (www.phoenixpeptide.com; 530 Harbor Boulevard, Belmont, CA), Peninsula Labs (San Carlos CA), SIGMA (St.Louis, MO www.sigma-aldrich.com), Cappel ICN (Irvine, California, www.icnbiomed.com), and Calbiochem (La Jolla, California, www.calbiochem.com), which are all incorporated herein by reference in their entirety. The polynucleotide sequences encoding these antibodies

may be obtained from the scientific literature, from patents, and from databases such as GenBank. Alternatively, one of ordinary skill in the art may design the polynucleotide sequence to be incorporated into the genome by choosing the codons that encode for each amino acid in the desired antibody. Antibodies made by the

5 transgenic animals of the present invention include antibodies that may be used as therapeutic reagents, for example in cancer immunotherapy against specific antigens, as diagnostic reagents and as laboratory reagents for numerous applications including immunoneutralization, radioimmunoassay, western blots, dot blots, ELISA, immunoprecipitation and immunoaffinity columns. Some of these antibodies include,

10 but are not limited to, antibodies which bind the following ligands: adrenomedullin, amylin, calcitonin, amyloid, calcitonin gene-related peptide, cholecystokinin, gastrin, gastric inhibitory peptide, gastrin releasing peptide, interleukin, interferon, cortistatin, somatostatin, endothelin, sarafotoxin, glucagon, glucagon-like peptide, insulin, atrial natriuretic peptide, BNP, CNP, neurokinin, substance P, leptin, neuropeptide Y,

15 melanin concentrating hormone, melanocyte stimulating hormone, orphanin, endorphin, dynorphin, enkephalin, enkephalin, leumorphin, peptide F, PACAP, PACAP-related peptide, parathyroid hormone, urocortin, corticotrophin releasing hormone, PHM, PHI, vasoactive intestinal polypeptide, secretin, ACTH, angiotensin, angiotensin, bombesin, endostatin, bradykinin, FMRF amide, galanin, gonadotropin

20 releasing hormone (GnRH) associated peptide, GnRH, growth hormone releasing hormone, inhibin, granulocyte-macrophage colony stimulating factor (GM-CSF), motilin, neurotensin, oxytocin, vasopressin, osteocalcin, pancreastatin, pancreatic polypeptide, peptide YY, proopiomelanocortin, transforming growth factor, vascular endothelial growth factor, vesicular monoamine transporter, vesicular acetylcholine

25 transporter, ghrelin, NPW, NPB, C3d, prokineticin, thyroid stimulating hormone, luteinizing hormone, follicle stimulating hormone, prolactin, growth hormone, beta-lipotropin, melatonin, kallikriens, kinins, prostaglandins, erythropoietin, p146 (SEQ ID NO:24 amino acid sequence, SEQ ID NO:25, nucleotide sequence), estrogen, testosterone, corticosteroids, mineralocorticoids, thyroid hormone, thymic hormones,

30 connective tissue proteins, nuclear proteins, actin, avidin, activin, agrin, albumin, and prohormones, propeptides, splice variants, fragments and analogs thereof.

The following is yet another non-limiting list of antibodies that can be produced by the methods of present invention: abciximab (ReoPro), abciximab anti-platelet aggregation monoclonal antibody, anti-CD11a (hu1124), anti-CD18 antibody,

anti-CD20 antibody, anti-cytomegalovirus (CMV) antibody, anti-digoxin antibody, anti-hepatitis B antibody, anti-HER-2 antibody, anti-idiotypic antibody to GD3 glycolipid, anti-IgE antibody, anti-IL-2R antibody, antimetastatic cancer antibody (mAb 17-1A), anti-rabies antibody, anti-respiratory syncytial virus (RSV) antibody, 5 anti-Rh antibody, anti-TCR, anti-TNF antibody, anti-VEGF antibody and fab fragment thereof, rattlesnake venom antibody, black widow spider venom antibody, coral snake venom antibody, antibody against very late antigen-4 (VLA-4), C225 humanized antibody to EGF receptor, chimeric (human & mouse) antibody against TNF α , antibody directed against GPIIb/IIIa receptor on human platelets, gamma 10 globulin, anti-hepatitis B immunoglobulin, human anti-D immunoglobulin, human antibodies against *S. aureus*, human tetanus immunoglobulin, humanized antibody against the epidermal growth receptor-2, humanized antibody against the α subunit of the interleukin-2 receptor, humanized antibody CTLA4IG, humanized antibody to the IL-2 R α -chain, humanized anti-CD40-ligand monoclonal antibody (5c8), humanized 15 mAb against the epidermal growth receptor-2, humanized mAb to rous sarcoma virus, humanized recombinant antibody (IgG1k) against respiratory syncytial virus (RSV), lymphocyte immunoglobulin (anti-thymocyte antibody), lymphocyte immunoglobulin, mAb against factor VII, MDX-210 bi-specific antibody against HER-2, MDX-22, MDX-220 bi-specific antibody against TAG-72 on tumors, MDX- 20 33 antibody to Fc γ R1 receptor, MDX-447 bi-specific antibody against EGF receptor, MDX-447 bispecific humanized antibody to EGF receptor, MDX-RA immunotoxin (ricin A linked) antibody, Medi-507 antibody (humanized form of BTL-322) against CD2 receptor on T-cells, monoclonal antibody LDP-02, muromonab-CD3(OKT3) antibody, OKT3 ("muromonab-CD3") antibody, PRO 542 antibody, ReoPro 25 ("abciximab") antibody, and TNF-IgG fusion protein.

The antibodies prepared using the methods of the present invention may also be designed to possess specific labels that may be detected through means known to one of ordinary skill in the art. The antibodies may also be designed to possess specific sequences useful for purification through means known to one of ordinary 30 skill in the art. Specialty antibodies designed for binding specific antigens may also be made in transgenic animals using the transposon-based vectors of the present invention.

Production of a monoclonal antibody using the transposon-based vectors of the present invention can be accomplished in a variety of ways. In one embodiment, two vectors may be constructed: one that encodes the light chain, and a second vector that encodes the heavy chain of the monoclonal antibody. These vectors may then be
5 incorporated into the genome of the target animal by methods disclosed herein. In an alternative embodiment, the sequences encoding light and heavy chains of a monoclonal antibody may be included on a single DNA construct. For example, the coding sequence of light and heavy chains of a murine monoclonal antibody that show specificity for human seminoprotein can be expressed using transposon-based
10 constructs of the present invention (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively).

Further included in the present invention are proteins and peptides synthesized by the immune system including those synthesized by the thymus, lymph nodes, spleen, and the gastrointestinal associated lymph tissues (GALT) system. The
15 immune system proteins and peptides proteins that can be made in transgenic animals using the transposon-based vectors of the present invention include, but are not limited to, alpha-interferon, beta-interferon, gamma-interferon, alpha-interferon A, alpha-interferon 1, G-CSF, GM-CSF, interleukin-1 (IL-1), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, TNF- α , and TNF- β . Other cytokines
20 included in the present invention include cardiotrophin, stromal cell derived factor, macrophage derived chemokine (MDC), melanoma growth stimulatory activity (MGSA), macrophage inflammatory proteins 1 alpha (MIP-1 alpha), 2, 3 alpha, 3 beta, 4 and 5.

Lytic peptides such as p146 are also included in the desired molecules of the
25 present invention. In one embodiment, the p146 peptide comprises an amino acid sequence of SEQ ID NO:24. The present invention also encompasses a transposon-based vector comprising a p146 nucleic acid comprising a polynucleotide sequence of SEQ ID NO:25.

Enzymes are another class of proteins that may be made through the use of the
30 transposon-based vectors of the present invention. Such enzymes include but are not limited to adenosine deaminase, alpha-galactosidase, cellulase, collagenase, dnaseI, hyaluronidase, lactase, L-asparaginase, pancreatin, papain, streptokinase B, subtilisin, superoxide dismutase, thrombin, trypsin, urokinase, fibrinolysin, glucocerebrosidase

and plasminogen activator. In some embodiments wherein the enzyme could have deleterious effects, additional amino acids and a protease cleavage site are added to the carboxy end of the enzyme of interest in order to prevent expression of a functional enzyme. Subsequent digestion of the enzyme with a protease results in
5 activation of the enzyme.

Extracellular matrix proteins are one class of desired proteins that may be made through the use of the present invention. Examples include but are not limited to collagen, fibrin, elastin, laminin, and fibronectin and subtypes thereof. Intracellular proteins and structural proteins are other classes of desired proteins in the present
10 invention.

Growth factors are another desired class of proteins that may be made through the use of the present invention and include, but are not limited to, transforming growth factor- α ("TGF- α "), transforming growth factor- β (TGF- β), platelet-derived growth factors (PDGF), fibroblast growth factors (FGF), including FGF acidic
15 isoforms 1 and 2, FGF basic form 2 and FGF 4, 8, 9 and 10, nerve growth factors (NGF) including NGF 2.5s, NGF 7.0s and beta NGF and neurotrophins, brain derived neurotrophic factor, cartilage derived factor, growth factors for stimulation of the production of red blood cells, growth factors for stimulation of the production of white blood cells, bone growth factors (BGF), basic fibroblast growth factor, vascular
20 endothelial growth factor (VEGF), granulocyte colony stimulating factor (G-CSF), insulin like growth factor (IGF) I and II, hepatocyte growth factor, glial neurotrophic growth factor (GDNF), stem cell factor (SCF), keratinocyte growth factor (KGF), transforming growth factors (TGF), including TGFs alpha, beta, beta1, beta2, beta3, skeletal growth factor, bone matrix derived growth factors, bone derived growth
25 factors, erythropoietin (EPO) and mixtures thereof.

Another desired class of proteins that may be made may be made through the use of the present invention include, but are not limited to, leptin, leukemia inhibitory factor (LIF), tumor necrosis factor alpha and beta, ENBREL, angiostatin, endostatin, thrombospondin, osteogenic protein-1, bone morphogenetic proteins 2 and 7,
30 osteonectin, somatomedin-like peptide, and osteocalcin.

Yet another desired class of proteins are blood proteins or clotting cascade protein including albumin, Prekallikrein, High molecular weight kininogen (HMWK) (contact activation cofactor; Fitzgerald, Flaujeac Williams factor), Factor I

(Fibrinogen), Factor II (prothrombin), Factor III (Tissue Factor), Factor IV (calcium), Factor V (proaccelerin, labile factor, accelerator (Ac-) globulin), Factor VI (Va) (accelerin), Factor VII (proconvertin), serum prothrombin conversion accelerator (SPCA), cothromboplastin), Factor VIII (antihemophilic factor A, antihemophilic
5 globulin (AHG)), Factor IX (Christmas Factor, antihemophilic factor B, plasma thromboplastin component (PTC)), Factor X (Stuart-Prower Factor), Factor XI (Plasma thromboplastin antecedent (PTA)), Factor XII (Hageman Factor), Factor XIII (rotransglutaminase, fibrin stabilizing factor (FSF), fibrinolygase), von Willebrand factor, Protein C, Protein S, Thrombomodulin, Antithrombin III.

10 A non-limiting list of the peptides and proteins that may be made may be made through the use of the present invention is provided in product catalogs of companies such as Phoenix Pharmaceuticals, Inc. (www.phoenixpeptide.com; 530 Harbor Boulevard, Belmont, CA), Peninsula Labs (San Carlos CA), SIGMA, (St.Louis, MO www.sigma-aldrich.com), Cappel ICN (Irvine, California,
15 www.icnbiomed.com), and Calbiochem (La Jolla, California, www.calbiochem.com). The polynucleotide sequences encoding these proteins and peptides of interest may be obtained from the scientific literature, from patents, and from databases such as GenBank. Alternatively, one of ordinary skill in the art may design the polynucleotide sequence to be incorporated into the genome by choosing the codons
20 that encode for each amino acid in the desired protein or peptide.

Some of these desired proteins or peptides that may be made through the use of the present invention include but are not limited to the following: adrenomedulin, amylin, calcitonin, amyloid, calcitonin gene-related peptide, cholecystokinin, gastrin, gastric inhibitory peptide, gastrin releasing peptide, interleukin, interferon, cortistatin,
25 somatostatin, endothelin, sarafotoxin, glucagon, glucagon-like peptide, insulin, atrial natriuretic peptide, BNP, CNP, neurokinin, substance P, leptin, neuropeptide Y, melanin concentrating hormone, melanocyte stimulating hormone, orphanin, endorphin, dynorphin, enkephalin, leumorphin, peptide F, PACAP, PACAP-related peptide, parathyroid hormone, urocortin, corticotrophin releasing hormone, PHM,
30 PHI, vasoactive intestinal polypeptide, secretin, ACTH, angiotensin, angiotatin, bombesin, endostatin, bradykinin, FMRF amide, galanin, gonadotropin releasing hormone (GnRH) associated peptide, GnRH, growth hormone releasing hormone, inhibin, granulocyte-macrophage colony stimulating factor (GM-CSF), motilin, neurotensin, oxytocin, vasopressin, osteocalcin, pancreastatin, pancreatic polypeptide,

- peptide YY, proopiomelanocortin, transforming growth factor, vascular endothelial growth factor, vesicular monoamine transporter, vesicular acetylcholine transporter, ghrelin, NPW, NPB, C3d, prokineticin, thyroid stimulating hormone, luteinizing hormone, follicle stimulating hormone, prolactin, growth hormone, beta-lipotropin, melatonin, kallikriens, kinins, prostaglandins, erythropoietin, p146 (SEQ ID NO:24, amino acid sequence, SEQ ID NO:25, nucleotide sequence), thymic hormones, connective tissue proteins, nuclear proteins, actin, avidin, activin, agrin, albumin, apolipoproteins, apolipoprotein A, apolipoprotein B, and prohormones, propeptides, splice variants, fragments and analogs thereof.
- Other desired proteins that may be made by the transgenic animals of the present invention include bacitracin, polymyxin b, vancomycin, cyclosporine, anti-RSV antibody, alpha-1 antitrypsin (AAT), anti-cytomegalovirus antibody, anti-hepatitis antibody, anti-inhibitor coagulant complex, anti-rabies antibody, anti-Rh(D) antibody, adenosine deaminase, anti-digoxin antibody, antivenin crotalidae (rattlesnake venom antibody), antivenin latrodectus (black widow spider venom antibody), antivenin micrurus (coral snake venom antibody), aprotinin, corticotropin (ACTH), diphtheria antitoxin, lymphocyte immune globulin (anti-thymocyte antibody), protamine, thyrotropin, capreomycin, α -galactosidase, gramicidin, streptokinase, tetanus toxoid, tyrothricin, IGF-1, proteins of varicella vaccine, anti-TNF antibody, anti-IL-2r antibody, anti-HER-2 antibody, OKT3 ("muromonab-CD3") antibody, TNF-IgG fusion protein, ReoPro ("abciximab") antibody, ACTH fragment 1-24, desmopressin, gonadotropin-releasing hormone, histrelin, leuprolide, lypressin, nafarelin, peptide that binds GPIIb/GPIIIa on platelets (integrilin), goserelin, capreomycin, colistin, anti-respiratory syncytial virus, lymphocyte immune globulin (Thymoglovin, Atgam), panorex, alpha-antitrypsin, botulinin, lung surfactant protein, tumor necrosis receptor-IgG fusion protein (enbrel), gonadorelin, proteins of influenza vaccine, proteins of rotavirus vaccine, proteins of haemophilus b conjugate vaccine, proteins of poliovirus vaccine, proteins of pneumococcal conjugate vaccine, proteins of meningococcal C vaccine, proteins of influenza vaccine, megakaryocyte growth and development factor (MGDF), neuroimmunophilin ligand-A (NIL-A), brain-derived neurotrophic factor (BDNF), glial cell line-derived neurotrophic factor (GDNF), leptin (native), leptin B, leptin C, IL-1RA (interleukin-1RA), R-568, novel erythropoiesis-stimulating protein (NESP), humanized mAb to rous sarcoma virus

(MEDI-493), glutamyl-tryptophan dipeptide IM862, LFA-3TIP immunosuppressive, humanized anti-CD40-ligand monoclonal antibody (5c8), gelsolin enzyme, tissue factor pathway inhibitor (TFPI), proteins of meningitis B vaccine, antimetastatic cancer antibody (mAb 17-1A), chimeric (human & mouse) mAb against TNF α , mAb
5 against factor VII, relaxin, capreomycin, glycopeptide (LY333328), recombinant human activated protein C (rhAPC), humanized mAb against the epidermal growth receptor-2, altepase, anti-CD20 antigen, C2B8 antibody, insulin-like growth factor-1, atrial natriuretic peptide (anaritide), tenectapase, anti-CD11a antibody (hu 1124), anti-CD18 antibody, mAb LDP-02, anti-VEGF antibody, fab fragment of anti-VEGF
10 Ab, APO2 ligand (tumor necrosis factor-related apoptosis-inducing ligand), rTGF- β (transforming growth factor- β), alpha-antitrypsin, ananain (a pineapple enzyme), humanized mAb CTLA4IG, PRO 542 (mAb), D2E7 (mAb), calf intestine alkaline phosphatase, α -L-iduronidase, α -L-galactosidase (humanglutamic acid decarboxylase, acid sphingomyelinase, bone morphogenetic protein-2 (rhBMP-2), proteins of HIV
15 vaccine, T cell receptor (TCR) peptide vaccine, TCR peptides, V beta 3 and V beta 13.1. (IR502), (IR501), BI 1050/1272 mAb against very late antigen-4 (VLA-4), C225 humanized mAb to EGF receptor, anti-idiotypic antibody to GD3 glycolipid, antibacterial peptide against *H. pylori*, MDX-447 bispecific humanized mAb to EGF receptor, anti-cytomegalovirus (CMV), Medi-491 B19 parvovirus vaccine, humanized
20 recombinant mAb (IgG1k) against respiratory syncytial virus (RSV), urinary tract infection vaccine (against "pili" on *Escherichia coli* strains), proteins of lyme disease vaccine against *B. burgdorferi* protein (DbpA), proteins of Medi-501 human papilloma virus-11 vaccine (HPV), *Streptococcus pneumoniae* vaccine, Medi-507 mAb (humanized form of BTL-322) against CD2 receptor on T-cells, MDX-33 mAb
25 to Fc γ R1 receptor, MDX-RA immunotoxin (ricin A linked) mAb, MDX-210 bi-specific mAb against HER-2, MDX-447 bi-specific mAb against EGF receptor, MDX-22, MDX-220 bi-specific mAb against TAG-72 on tumors, colony-stimulating factor (CSF) (molgramostim), humanized mAb to the IL-2 R α -chain (basiliximab), mAb to IgE (IGE 025A), myelin basic protein-altered peptide (MSP771A),
30 humanized mAb against the epidermal growth receptor-2, humanized mAb against the α subunit of the interleukin-2 receptor, low molecular weight heparin, anti-hemophillic factor, and bactericidal/permeability-increasing protein (r-BPI).

The peptides and proteins made using the present invention may be labeled using labels and techniques known to one of ordinary skill in the art. Some of these

labels are described in the "Handbook of Fluorescent Probes and Research Products", ninth edition, Richard P. Haugland (ed) Molecular Probes, Inc. Eugene, OR), which is incorporated herein in its entirety. Some of these labels may be genetically engineered into the polynucleotide sequence for the expression of the selected protein or peptide. The peptides and proteins may also have label-incorporation "handles" incorporated to allow labeling of an otherwise difficult or impossible to label protein.

It is to be understood that the various classes of desired peptides and proteins, as well as specific peptides and proteins described in this section may be modified as described below by inserting selected codons for desired amino acid substitutions into the gene incorporated into the transgenic animal.

The present invention may also be used to produce desired molecules other than proteins and peptides including, but not limited to, lipoproteins such as high density lipoprotein (HDL), HDL-Milano, and low density lipoprotein, lipids, carbohydrates, siRNA and ribozymes. In these embodiments, a gene of interest encodes a nucleic acid molecule or a protein that directs production of the desired molecule.

The present invention further encompasses the use of inhibitory molecules to inhibit endogenous (i.e., non-vector) protein production. These inhibitory molecules include antisense nucleic acids, siRNA and inhibitory proteins. In a preferred embodiment, the endogenous protein whose expression is inhibited is an egg white protein including, but not limited to ovalbumin, ovotransferrin, and ovomucin. In one embodiment, a transposon-based vector containing an ovalbumin DNA sequence, that upon transcription forms a double stranded RNA molecule, is transfected into an animal such as a bird and the bird's production of endogenous ovalbumin protein is reduced by the interference RNA mechanism (RNAi). In other embodiments, a transposon-based vector encodes an inhibitory RNA molecule that inhibits the expression of more than one egg white protein. One exemplary construct is provided in Figure 4 wherein "Ovgen" indicates approximately 60 base pairs of an ovalbumin gene, "Ovotrans" indicates approximately 60 base pairs of an ovotransferrin gene and "Ovomucin" indicates approximately 60 base pairs of an ovomucin gene. These ovalbumin, ovotransferrin and ovomucin can be from any avian species, and in some embodiments, are from a chicken or quail. The term "pro" indicates the pro portion of a prepro sequence. One exemplary prepro sequence is that of cecropin and comprising base pairs 563-733 of the Cecropin cap site and Prepro provided in

- Genbank accession number X07404. Additional cecropin prepro and pro sequences are provided in SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, and SEQ ID NO:51. Additionally, inducible knockouts or knockdowns of the endogenous protein may be created to achieve a reduction or inhibition of endogenous protein production.
- 5 Endogenous egg white production can be inhibited in an avian at any time, but is preferably inhibited preceding, or immediately preceding, the harvest of eggs.

Modified Desired Proteins and Peptides

- "Proteins", "peptides," "polypeptides" and "oligopeptides" are chains of amino acids (typically L-amino acids) whose alpha carbons are linked through peptide bonds
- 10 formed by a condensation reaction between the carboxyl group of the alpha carbon of one amino acid and the amino group of the alpha carbon of another amino acid. The terminal amino acid at one end of the chain (i.e., the amino terminal) has a free amino group, while the terminal amino acid at the other end of the chain (i.e., the carboxy terminal) has a free carboxyl group. As such, the term "amino terminus" (abbreviated
- 15 N-terminus) refers to the free alpha-amino group on the amino acid at the amino terminal of the protein, or to the alpha-amino group (imino group when participating in a peptide bond) of an amino acid at any other location within the protein. Similarly, the term "carboxy terminus" (abbreviated C-terminus) refers to the free carboxyl group on the amino acid at the carboxy terminus of a protein, or to the
- 20 carboxyl group of an amino acid at any other location within the protein.

- Typically, the amino acids making up a protein are numbered in order, starting at the amino terminal and increasing in the direction toward the carboxy terminal of the protein. Thus, when one amino acid is said to "follow" another, that amino acid is positioned closer to the carboxy terminal of the protein than the preceding amino acid.

- 25 The term "residue" is used herein to refer to an amino acid (D or L) or an amino acid mimetic that is incorporated into a protein by an amide bond. As such, the amino acid may be a naturally occurring amino acid or, unless otherwise limited, may encompass known analogs of natural amino acids that function in a manner similar to the naturally occurring amino acids (i.e., amino acid mimetics). Moreover, an amide
- 30 bond mimetic includes peptide backbone modifications well known to those skilled in the art.

Furthermore, one of skill will recognize that, as mentioned above, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than about 5%, more typically less

- than about 1%) in an encoded sequence are conservatively modified variations where the alterations result in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. The following six groups each contain amino acids
- 5 that are conservative substitutions for one another:
- 1) Alanine (A), Serine (S), Threonine (T);
 - 2) Aspartic acid (D), Glutamic acid (E);
 - 3) Asparagine (N), Glutamine (Q);
 - 4) Arginine (R), Lysine (K);
 - 10 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and
 - 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

A conservative substitution is a substitution in which the substituting amino acid (naturally occurring or modified) is structurally related to the amino acid being substituted, i.e., has about the same size and electronic properties as the amino acid

15 being substituted. Thus, the substituting amino acid would have the same or a similar functional group in the side chain as the original amino acid. A "conservative substitution" also refers to utilizing a substituting amino acid which is identical to the amino acid being substituted except that a functional group in the side chain is protected with a suitable protecting group.

- 20 Suitable protecting groups are described in Green and Wuts, "Protecting Groups in Organic Synthesis", John Wiley and Sons, Chapters 5 and 7, 1991, the teachings of which are incorporated herein by reference. Preferred protecting groups are those which facilitate transport of the peptide through membranes, for example, by reducing the hydrophilicity and increasing the lipophilicity of the peptide, and which
- 25 can be cleaved, either by hydrolysis or enzymatically (Ditter et al., 1968. J. Pharm. Sci. 57:783; Ditter et al., 1968. J. Pharm. Sci. 57:828; Ditter et al., 1969. J. Pharm. Sci. 58:557; King et al., 1987. Biochemistry 26:2294; Lindberg et al., 1989. Drug Metabolism and Disposition 17:311; Tunek et al., 1988. Biochem. Pharm. 37:3867; Anderson et al., 1985 Arch. Biochem. Biophys. 239:538; and Singhal et al., 1987. FASEB J. 1:220). Suitable hydroxyl protecting groups include ester, carbonate and carbamate protecting groups. Suitable amine protecting groups include acyl groups and alkoxy or aryloxy carbonyl groups, as described above for N-terminal protecting groups. Suitable carboxylic acid protecting groups include aliphatic, benzyl and aryl esters, as described below for C-terminal protecting groups. In one embodiment, the
- 30

carboxylic acid group in the side chain of one or more glutamic acid or aspartic acid residues in a peptide of the present invention is protected, preferably as a methyl, ethyl, benzyl or substituted benzyl ester, more preferably as a benzyl ester.

- Provided below are groups of naturally occurring and modified amino acids in which each amino acid in a group has similar electronic and steric properties. Thus, a conservative substitution can be made by substituting an amino acid with another amino acid from the same group. It is to be understood that these groups are non-limiting, i.e. that there are additional modified amino acids which could be included in each group.
- 10 Group I includes leucine, isoleucine, valine, methionine and modified amino acids having the following side chains: ethyl, n-propyl n-butyl. Preferably, Group I includes leucine, isoleucine, valine and methionine.
- Group II includes glycine, alanine, valine and a modified amino acid having an ethyl side chain. Preferably, Group II includes glycine and alanine.
- 15 Group III includes phenylalanine, phenylglycine, tyrosine, tryptophan, cyclohexylmethyl glycine, and modified amino residues having substituted benzyl or phenyl side chains. Preferred substituents include one or more of the following: halogen, methyl, ethyl, nitro, $-\text{NH}_2$, methoxy, ethoxy and $-\text{CN}$. Preferably, Group III includes phenylalanine, tyrosine and tryptophan.
- 20 Group IV includes glutamic acid, aspartic acid, a substituted or unsubstituted aliphatic, aromatic or benzylic ester of glutamic or aspartic acid (e.g., methyl, ethyl, n-propyl iso-propyl, cyclohexyl, benzyl or substituted benzyl), glutamine, asparagine, $-\text{CO}-\text{NH}-$ alkylated glutamine or asparagines (e.g., methyl, ethyl, n-propyl and iso-propyl) and modified amino acids having the side chain $-(\text{CH}_2)_3-\text{COOH}$, an ester thereof (substituted or unsubstituted
- 25 aliphatic, aromatic or benzylic ester), an amide thereof and a substituted or unsubstituted N-alkylated amide thereof. Preferably, Group IV includes glutamic acid, aspartic acid, methyl aspartate, ethyl aspartate, benzyl aspartate and methyl glutamate, ethyl glutamate and benzyl glutamate, glutamine and asparagine.
- 30 Group V includes histidine, lysine, ornithine, arginine, N-nitroarginine, β -cycloarginine, γ -hydroxyarginine, N-amidinocitruline and 2-amino-4-guanidinobutanoic acid, homologs of lysine, homologs of arginine and homologs of ornithine. Preferably, Group V includes histidine, lysine,

arginine and ornithine. A homolog of an amino acid includes from 1 to about 3 additional or subtracted methylene units in the side chain.

Group VI includes serine, threonine, cysteine and modified amino acids having C1-C5 straight or branched alkyl side chains substituted with —OH or —SH, for example, —CH₂CH₂OH, —CH₂CH₂CH₂OH or —CH₂CH₂OHCH₃. Preferably, Group VI includes serine, cysteine or threonine.

In another aspect, suitable substitutions for amino acid residues include “severe” substitutions. A “severe substitution” is a substitution in which the substituting amino acid (naturally occurring or modified) has significantly different size and/or electronic properties compared with the amino acid being substituted. Thus, the side chain of the substituting amino acid can be significantly larger (or smaller) than the side chain of the amino acid being substituted and/or can have functional groups with significantly different electronic properties than the amino acid being substituted. Examples of severe substitutions of this type include the substitution of phenylalanine or cyclohexylmethyl glycine for alanine, isoleucine for glycine, a D amino acid for the corresponding L amino acid, or —NH—CH[(—CH₂)₅—COOH]—CO— for aspartic acid. Alternatively, a functional group may be added to the side chain, deleted from the side chain or exchanged with another functional group. Examples of severe substitutions of this type include adding of valine, leucine or isoleucine, exchanging the carboxylic acid in the side chain of aspartic acid or glutamic acid with an amine, or deleting the amine group in the side chain of lysine or ornithine. In yet another alternative, the side chain of the substituting amino acid can have significantly different steric and electronic properties than the functional group of the amino acid being substituted. Examples of such modifications include tryptophan for glycine, lysine for aspartic acid and —(CH₂)₄COOH for the side chain of serine. These examples are not meant to be limiting.

In another embodiment, for example in the synthesis of a peptide 26 amino acids in length, the individual amino acids may be substituted according in the following manner:

AA₁ is serine, glycine, alanine, cysteine or threonine;

AA₂ is alanine, threonine, glycine, cysteine or serine;

- AA₃ is valine, arginine, leucine, isoleucine, methionine, ornithine, lysine, N-nitroarginine, β-cycloarginine, γ-hydroxyarginine, N-amidinocitruline or 2-amino-4-guanidinobutanoic acid;
- AA₄ is proline, leucine, valine, isoleucine or methionine;
- 5 AA₅ is tryptophan, alanine, phenylalanine, tyrosine or glycine;
- AA₆ is serine, glycine, alanine, cysteine or threonine;
- AA₇ is proline, leucine, valine, isoleucine or methionine;
- AA₈ is alanine, threonine, glycine, cysteine or serine;
- AA₉ is alanine, threonine, glycine, cysteine or serine;
- 10 AA₁₀ is leucine, isoleucine, methionine or valine;
- AA₁₁ is serine, glycine, alanine, cysteine or threonine;
- AA₁₂ is leucine, isoleucine, methionine or valine;
- AA₁₃ is leucine, isoleucine, methionine or valine;
- AA₁₄ is glutamine, glutamic acid, aspartic acid, asparagine, or a substituted or
- 15 unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;
- AA₁₅ is arginine, N-nitroarginine, β-cycloarginine, γ-hydroxy-arginine, N-amidinocitruline or 2-amino-4-guanidino-butanoic acid
- AA₁₆ is proline, leucine, valine, isoleucine or methionine;
- AA₁₇ is serine, glycine, alanine, cysteine or threonine;
- 20 AA₁₈ is glutamic acid, aspartic acid, asparagine, glutamine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;
- AA₁₉ is aspartic acid, asparagine, glutamic acid, glutamine, leucine, valine, isoleucine, methionine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;
- 25 AA₂₀ is valine, arginine, leucine, isoleucine, methionine, ornithine, lysine, N-nitroarginine, β-cycloarginine, γ-hydroxyarginine, N-amidinocitruline or 2-amino-4-guanidinobutanoic acid;
- AA₂₁ is alanine, threonine, glycine, cysteine or serine;
- AA₂₂ is alanine, threonine, glycine, cysteine or serine;
- 30 AA₂₃ is histidine, serine, threonine, cysteine, lysine or ornithine;
- AA₂₄ is threonine, aspartic acid, serine, glutamic acid or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;

AA₂₅ is asparagine, aspartic acid,, glutamic acid, glutamine, leucine, valine, isoleucine, methionine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid; and

AA₂₆ is cysteine, histidine, serine, threonine, lysine or ornithine.

- 5 It is to be understood that these amino acid substitutions may be made for longer or shorter peptides than the 26 mer in the preceding example above, and for proteins.

 In one embodiment of the present invention, codons for the first several N-terminal amino acids of the transposase are modified such that the third base of each
10 codon is changed to an A or a T without changing the corresponding amino acid. It is preferable that between approximately 1 and 20, more preferably 3 and 15, and most preferably between 4 and 12 of the first N-terminal codons of the gene of interest are modified such that the third base of each codon is changed to an A or a T without changing the corresponding amino acid. In one embodiment, the first ten N-terminal
15 codons of the gene of interest are modified in this manner.

 When several desired proteins, protein fragments or peptides are encoded in the gene of interest to be incorporated into the genome, one of skill in the art will appreciate that the proteins, protein fragments or peptides may be separated by a
20 spacer molecule such as, for example, a peptide, consisting of one or more amino acids. Generally, the spacer will have no specific biological activity other than to join the desired proteins, protein fragments or peptides together, or to preserve some minimum distance or other spatial relationship between them. However, the constituent amino acids of the spacer may be selected to influence some property of the molecule such as the folding, net charge, or hydrophobicity. The spacer may also
25 be contained within a nucleotide sequence with a purification handle or be flanked by cleavage sites, such as proteolytic cleavage sites.

 Such polypeptide spacers may have from about 5 to about 40 amino acid residues. The spacers in a polypeptide are independently chosen, but are preferably all the same. The spacers should allow for flexibility of movement in space and are
30 therefore typically rich in small amino acids, for example, glycine, serine, proline or alanine. Preferably, peptide spacers contain at least 60%, more preferably at least 80% glycine or alanine. In addition, peptide spacers generally have little or no biological and antigenic activity. Preferred spacers are (Gly-Pro-Gly-Gly)_x (SEQ ID

NO:26) and (Gly₄-Ser)_y, wherein x is an integer from about 3 to about 9 and y is an integer from about 1 to about 8. Specific examples of suitable spacers include (Gly-Pro-Gly-Gly)₃

SEQ ID NO:27 Gly Pro Gly Gly Gly Pro Gly Gly Gly Pro Gly Gly

5 (Gly₄-Ser)₃

SEQ ID NO:28 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
or (Gly₄-Ser)₄

SEQ ID NO:29 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
Gly Gly Gly Gly Ser.

10 Nucleotide sequences encoding for the production of residues which may be useful in purification of the expressed recombinant protein may also be built into the vector. Such sequences are known in the art and include the glutathione binding domain from glutathione S-transferase, polylysine, hexa-histidine or other cationic amino acids, thioredoxin, hemagglutinin antigen and maltose binding protein.

15 Additionally, nucleotide sequences may be inserted into the gene of interest to be incorporated so that the protein or peptide can also include from one to about six amino acids that create signals for proteolytic cleavage. In this manner, if a gene is designed to make one or more peptides or proteins of interest in the transgenic animal, specific nucleotide sequences encoding for amino acids recognized by enzymes may
20 be incorporated into the gene to facilitate cleavage of the large protein or peptide sequence into desired peptides or proteins or both. For example, nucleotides encoding a proteolytic cleavage site can be introduced into the gene of interest so that a signal sequence can be cleaved from a protein or peptide encoded by the gene of interest. Nucleotide sequences encoding other amino acid sequences which display pH
25 sensitivity or chemical sensitivity may also be added to the vector to facilitate separation of the signal sequence from the peptide or protein of interest.

Proteolytic cleavage sites include cleavage sites recognized by exopeptidases such as carboxypeptidase A, carboxypeptidase B, aminopeptidase I, and dipeptidylaminopeptidase; endopeptidases such as trypsin, V8-protease, enterokinase,
30 factor Xa, collagenase, endoproteinase, subtilisin, and thrombin; and proteases such as Protease 3C IgA protease (Igase) Rhinovirus 3C(preScission)protease. Chemical cleavage sites are also included in the definition of cleavage site as used herein. Chemical cleavage sites include, but are not limited to, site cleaved by cyanogen bromide, hydroxylamine, formic acid, and acetic acid.

In one embodiment of the present invention, a TAG sequence is linked to the gene of interest. The TAG sequence serves three purposes: 1) it allows free rotation of the peptide or protein to be isolated so there is no interference from the native protein or signal sequence, i.e. vitellogenin, 2) it provides a "purification handle" to isolate the protein using column purification, and 3) it includes a cleavage site to remove the desired protein from the signal and purification sequences. Accordingly, as used herein, a TAG sequence includes a spacer sequence, a purification handle and a cleavage site. The spacer sequences in the TAG proteins contain one or more repeats shown in SEQ ID NO:30. A preferred spacer sequence comprises the sequence provided in SEQ ID NO:31. One example of a purification handle is the gp41 hairpin loop from HIV I. Exemplary gp41 polynucleotide and polypeptide sequences are provided in SEQ ID NO:32 and SEQ ID NO:33, respectively. However, it should be understood that any antigenic region may be used as a purification handle, including any antigenic region of gp41. Preferred purification handles are those that elicit highly specific antibodies. Additionally, the cleavage site can be any protein cleavage site known to one of ordinary skill in the art and includes an enterokinase cleavage site comprising the Asp Asp Asp Asp Lys sequence (SEQ ID NO:34) and a furin cleavage site. Constructs containing a TAG sequence are shown in Figures 2 and 3. In one embodiment of the present invention, the TAG sequence comprises a polynucleotide sequence of SEQ ID NO:35.

Methods of Administering Transposon-Based Vectors

In addition to the transposon-based vectors described above, the present invention also includes methods of administering the transposon-based vectors to an animal, methods of producing a transgenic animal wherein a gene of interest is incorporated into the germline of the animal and methods of producing a transgenic animal wherein a gene of interest is incorporated into cells other than the germline cells (somatic cells) of the animal. The transposon-based vectors of the present invention are administered to a reproductive organ of an animal via any method known to those of skill in the art. Preferred reproductive organs include an ovary, an oviduct, a mammary gland, and a fallopian tube.

In some embodiments, a transposon-based vector is directly administered to the reproductive organ. Direct administration encompasses injection into the organ, and in a preferred embodiment, a transposon-based vector is injected into the lumen of the oviduct, and more preferably, the lumen of the magnum or the infundibulum of

the oviduct. The transposon-based vectors may additionally or alternatively be placed in an artery supplying the reproductive organ. Administering the vectors to the artery supplying the ovary results in transfection of follicles and oocytes in the ovary to create a germline transgenic animal. Alternatively, supplying the vectors through an artery leading to the oviduct would preferably transfect the tubular gland and epithelial cells. Such transfected cells could manufacture a desired protein or peptide for deposition in the egg white. In one embodiment, a transposon-based vector is administered into the lumen of the magnum or the infundibulum of the oviduct and to an artery supplying the oviduct. Indirect administration to the oviduct epithelium may occur through the cloaca. Direct administration into the mammary gland comprises introduction into the duct system of the mammary gland.

Administration of transposon-based vectors may occur in arteries supplying the ovary and or through direct intrathecal administration into the ovary through injection.

The transposon-based vectors may be administered in a single administration, multiple administrations, continuously, or intermittently. The transposon-based vectors may be administered by injection, via a catheter, an osmotic mini-pump or any other method. In some embodiments, the transposon-based vector is administered to an animal in multiple administrations, each administration containing the vector and a different transfecting reagent.

The transposon-based vectors may be administered to the animal at any point during the lifetime of the animal, however, it is preferable that the vectors are administered prior to the animal reaching sexual maturity. The transposon-based vectors are preferably administered to a chicken between approximately 14 and 16 weeks of age and to a quail between approximately 5 and 10 weeks of age, more preferably 5 and 8 weeks of age, and most preferably between 5 and 6 weeks of age, when standard poultry rearing practices are used. The vectors may be administered at earlier ages when exogenous hormones are used to induce early sexual maturation in the bird. In some embodiments, the transposon-based vector is administered to an animal following an increase in proliferation of the oviduct epithelial cells and/or the tubular gland cells. Such an increase in proliferation normally follows an influx of reproductive hormones in the area of the oviduct. When the animal is an avian, the transposon-based vector is administered following an increase in proliferation of the oviduct epithelial cells and before the avian begins to produce egg white constituents.

In a preferred embodiment, the animal is an egg-laying animal, and more preferably, an avian. In one embodiment, between approximately 1 and 150 μg , 1 and 100 μg , 1 and 50 μg , preferably between 1 and 20 μg , and more preferably between 5 and 10 μg of transposon-based vector DNA is administered to the oviduct of a bird.

5 Optimal ranges depend upon the type of bird and the bird's stage of sexual maturity. In a chicken, it is preferred that between approximately 1 and 100 μg , or 5 and 50 μg are administered. In a quail, it is preferred that between approximately 5 and 10 μg are administered. Intraoviduct administration of the transposon-based vectors of the present invention result in incorporation of the gene of interest into the cells of the

10 oviduct as evidenced by a PCR positive signal in the oviduct tissue. In other embodiments, the transposon-based vector is administered to an artery that supplies the oviduct. These methods of administration may also be combined with any methods for facilitating transfection, including without limitation, electroporation, gene guns, injection of naked DNA, and use of dimethyl sulfoxide (DMSO).

15 According to the present invention, the transposon-based vector is administered in conjunction with an acceptable carrier and/or transfection reagent. Acceptable carriers include, but are not limited to, water, saline, Hanks Balanced Salt Solution (HBSS), Tris-EDTA (TE) and lyotropic liquid crystals. Transfection reagents commonly known to one of ordinary skill in the art that may be employed

20 include, but are not limited to, the following: cationic lipid transfection reagents, cationic lipid mixtures, polyamine reagents, liposomes and combinations thereof; SUPERFECT®, Cytofectene, BioPORTER®, GenePORTER®, NeuroPORTER®, and perfectin from Gene Therapy Systems; lipofectamine, cellfectin, DMRIE-C oligofectamine, TROJENE® and PLUS reagent from InVitrogen; Xtreme gene, fugene, DOSPER and DOTAP from Roche; Lipotaxi and Genejammer from Strategene; and Escort from SIGMA. In one embodiment, the transfection reagent is SUPERFECT®. The ratio of DNA to transfection reagent may vary based upon the method of administration. In one embodiment, the transposon-based vector is

25 administered to the oviduct and the ratio of DNA to transfection reagent can be from 1:1.5 to 1:15, preferably 1:2 to 1:5, all expressed as wt/vol. Transfection may also be

30 accomplished using other means known to one of ordinary skill in the art, including without limitation electroporation, gene guns, injection of naked DNA, and use of dimethyl sulfoxide (DMSO).

Depending upon the cell or tissue type targeted for transfection, the form of the transposon-based vector may be important. Plasmids harvested from bacteria are generally closed circular supercoiled molecules, and this is the preferred state of a vector for gene delivery because of the ease of preparation. In some instances, transposase expression and insertion may be more efficient in a relaxed, closed circular configuration or in a linear configuration. In still other instances, a purified transposase protein may be co-injected with a transposon-based vector containing the gene of interest for more immediate insertion. This could be accomplished by using a transfection reagent complexed with both the purified transposase protein and the transposon-based vector.

Testing for and Breeding Animals Carrying the Transgene

Following administration of a transposon-based vector to an animal, DNA is extracted from the animal to confirm integration of the gene of interest. Advantages provided by the present invention include the high rates of integration, or incorporation, and transcription of the gene of interest when administered to a bird via an intraoviduct or intraovarian route (including intraarterial administrations to arteries leading to the oviduct or ovary). Example 6 below describes isolation of a proinsulin/ENT TAG protein from a transgenic hen following ammonium sulfate precipitation and ion exchange chromatography. Figure 5 demonstrates successful administration of a transposon-based vector to a hen, successful integration of the gene of interest, successful production of a protein encoded by the gene of interest, and successful deposition of the protein in egg white produced by the transgenic hen.

Actual frequencies of integration may be estimated both by comparative strength of the PCR signal, and by histological evaluation of the tissues by quantitative PCR. Another method for estimating the rate of transgene insertion is the so-called primed in situ hybridization technique (PRINS). This method determines not only which cells carry a transgene of interest, but also into which chromosome the gene has inserted, and even what portion of the chromosome. Briefly, labeled primers are annealed to chromosome spreads (affixed to glass slides) through one round of PCR, and the slides are then developed through normal in situ hybridization procedures. This technique combines the best features of in situ PCR and fluorescence in situ hybridization (FISH) to provide distinct chromosome location and copy number of the gene in question.

Breeding experiments are also conducted to determine if germline transmission of the transgene has occurred. In a general bird breeding experiment performed according to the present invention, each male bird was exposed to 2-3 different adult female birds for 3-4 days each. This procedure was continued with 5 different females for a total period of 6-12 weeks. Eggs are collected daily for up to 14 days after the last exposure to the transgenic male, and each egg is incubated in a standard incubator. The resulting embryos are examined for transgene presence at day 3 or 4 using PCR. It is to be understood that the above procedure can be modified to suit animals other than birds and that selective breeding techniques may be 10 performed to amplify gene copy numbers and protein output.

Production of Desired Proteins or Peptides in Egg White

In one embodiment, the transposon-based vectors of the present invention may be administered to a bird for production of desired proteins or peptides in the egg white. These transposon-based vectors preferably contain one or more of an 15 ovalbumin promoter, an ovomucoid promoter, an ovalbumin signal sequence and an ovomucoid signal sequence. Oviduct-specific ovalbumin promoters are described in B. O'Malley et al., 1987. EMBO J., vol. 6, pp. 2305-12; A. Qiu et al., 1994. Proc. Nat. Acad. Sci. (USA), vol. 91, pp. 4451-4455; D. Monroe et al., 2000. Biochim. Biophys. Acta, 1517 (1):27-32; H. Park et al., 2000. Biochem., 39:8537-8545; and T. 20 Muramatsu et al., 1996. Poul. Avian Biol. Rev., 6:107-123. Examples of transposon-based vectors designed for production of a desired protein in an egg white are shown in Figures 2 and 3.

Production of Desired Proteins or Peptides in Egg Yolk

The present invention is particularly advantageous for production of 25 recombinant peptides and proteins of low solubility in the egg yolk. Such proteins include, but are not limited to, membrane-associated or membrane-bound proteins, lipophilic compounds; attachment factors, receptors, and components of second messenger transduction machinery. Low solubility peptides and proteins are particularly challenging to produce using conventional recombinant protein 30 production techniques (cell and tissue cultures) because they aggregate in water-based, hydrophilic environments. Such aggregation necessitates denaturation and refolding of the recombinantly-produced proteins, which may deleteriously affect their structure and function. Moreover, even highly soluble recombinant peptides and proteins may precipitate and require denaturation and renaturation when produced in

sufficiently high amounts in recombinant protein production systems. The present invention provides an advantageous resolution of the problem of protein and peptide solubility during production of large amounts of recombinant proteins.

In one embodiment of the present invention wherein germline transfection is obtained via intraovarian administration of the transposon-based vector, deposition of a desired protein into the egg yolk is accomplished in offspring by attaching a sequence encoding a protein capable of binding to the yolk vitellogenin receptor to a gene of interest that encodes a desired protein. This transposon-based vector can be used for the receptor-mediated uptake of the desired protein by the oocytes. In a preferred embodiment, the sequence ensuring the binding to the vitellogenin receptor is a targeting sequence of a vitellogenin protein. The invention encompasses various vitellogenin proteins and their targeting sequences. In a preferred embodiment, a chicken vitellogenin protein targeting sequence is used, however, due to the high degree of conservation among vitellogenin protein sequences and known cross-species reactivity of vitellogenin targeting sequences with their egg-yolk receptors, other vitellogenin targeting sequences can be substituted. One example of a construct for use in the transposon-based vectors of the present invention and for deposition of an insulin protein in an egg yolk is a transposon-based vector containing a vitellogenin promoter, a vitellogenin targeting sequence, a TAG sequence, a pro-insulin sequence and a synthetic polyA sequence. The present invention includes, but is not limited to, vitellogenin targeting sequences residing in the N-terminal domain of vitellogenin, particularly in lipovitellin I. In one embodiment, the vitellogenin targeting sequence contains the polynucleotide sequence of SEQ ID NO:22. In a preferred embodiment, the transposon-based vector contains a transposase gene operably-linked to a constitutive promoter and a gene of interest operably-linked to a liver-specific promoter and a vitellogenin targeting sequence.

Isolation and Purification of Desired Protein or Peptide

For large-scale production of protein, an animal breeding stock that is homozygous for the transgene is preferred. Such homozygous individuals are obtained and identified through, for example, standard animal breeding procedures or PCR protocols.

Once expressed, peptides, polypeptides and proteins can be purified according to standard procedures known to one of ordinary skill in the art, including ammonium sulfate precipitation, affinity columns, column chromatography, gel electrophoresis,

high performance liquid chromatography, immunoprecipitation and the like. Substantially pure compositions of about 50 to 99% homogeneity are preferred, and 80 to 95% or greater homogeneity are most preferred for use as therapeutic agents.

In one embodiment of the present invention, the animal in which the desired protein is produced is an egg-laying animal. In a preferred embodiment of the present invention, the animal is an avian and a desired peptide, polypeptide or protein is isolated from an egg white. Egg white containing the exogenous protein or peptide is separated from the yolk and other egg constituents on an industrial scale by any of a variety of methods known in the egg industry. See, e.g., W. Stadelman et al. (Eds.), *Egg Science & Technology*, Haworth Press, Binghamton, NY (1995). Isolation of the exogenous peptide or protein from the other egg white constituents is accomplished by any of a number of polypeptide isolation and purification methods well known to one of ordinary skill in the art. These techniques include, for example, chromatographic methods such as gel permeation, ion exchange, affinity separation, metal chelation, HPLC, and the like, either alone or in combination. Another means that may be used for isolation or purification, either in lieu of or in addition to chromatographic separation methods, includes electrophoresis. Successful isolation and purification is confirmed by standard analytic techniques, including HPLC, mass spectroscopy, and spectrophotometry. These separation methods are often facilitated if the first step in the separation is the removal of the endogenous ovalbumin fraction of egg white, as doing so will reduce the total protein content to be further purified by about 50%.

To facilitate or enable purification of a desired protein or peptide, transposon-based vectors may include one or more additional epitopes or domains. Such epitopes or domains include DNA sequences encoding enzymatic or chemical cleavage sites including, but not limited to, an enterokinase cleavage site; the glutathione binding domain from glutathione S-transferase; polylysine; hexa-histidine or other cationic amino acids; thioredoxin; hemagglutinin antigen; maltose binding protein; a fragment of gp41 from HIV; and other purification epitopes or domains commonly known to one of skill in the art.

In one representative embodiment, purification of desired proteins from egg white utilizes the antigenicity of the ovalbumin carrier protein and particular attributes of a TAG linker sequence that spans ovalbumin and the desired protein. The TAG sequence is particularly useful in this process because it contains 1) a highly antigenic

epitope, a fragment of gp41 from HIV, allowing for stringent affinity purification, and, 2) a recognition site for the protease enterokinase immediately juxtaposed to the desired protein. In a preferred embodiment, the TAG sequence comprises approximately 50 amino acids. A representative TAG sequence is provided below.

5

Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp
Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Thr Thr Cys Ile Leu Lys Gly Ser Cys
Gly Trp Ile Gly Leu Leu Asp Asp Asp Asp Lys (SEQ ID NO:35)

- 10 The underlined sequences were taken from the hairpin loop domain of HIV gp-41 (SEQ ID NO:33). Sequences in italics represent the cleavage site for enterokinase (SEQ ID NO:34). The spacer sequence upstream of the loop domain was made from repeats of (Pro Ala Asp Asp Ala) (SEQ ID NO:31) to provide free rotation and promote surface availability of the hairpin loop from the ovalbumin carrier protein.

- 15 Isolation and purification of a desired protein is performed as follows:

1. Enrichment of the egg white protein fraction containing ovalbumin and the transgenic ovalbumin-TAG-desired protein.
2. Size exclusion chromatography to isolate only those proteins within a narrow range of molecular weights (a further enrichment of step 1).
- 20 3. Ovalbumin affinity chromatography. Highly specific antibodies to ovalbumin will eliminate virtually all extraneous egg white proteins except ovalbumin and the transgenic ovalbumin-TAG-desired protein.
4. gp41 affinity chromatography using anti-gp41 antibodies. Stringent application of this step will result in virtually pure transgenic ovalbumin-TAG-desired protein.
- 25 5. Cleavage of the transgene product can be accomplished in at least one of two ways:
 - a. The transgenic ovalbumin-TAG-desired protein is left attached to the gp41 affinity resin (beads) from step 4 and the protease enterokinase is added. This liberates the transgene target protein from the gp41 affinity resin while the ovalbumin-TAG sequence is retained. Separation by centrifugation (in a batch process) or flow through (in a column purification), leaves the desired protein together with enterokinase in solution. Enterokinase is recovered and reused.
- 30

- b. Alternatively, enterokinase is immobilized on resin (beads) by the addition of poly-lysine moieties to a non-catalytic area of the protease. The transgenic ovalbumin-TAG-desired protein eluted from the affinity column of step 4 is then applied to the protease resin. Protease action cleaves the ovalbumin-TAG sequence from the desired protein and leaves both entities in solution. The immobilized enterokinase resin is recharged and reused.
- c. The choice of these alternatives is made depending upon the size and chemical composition of the transgene target protein.

6. A final separation of either of these two (5a or 5b) protein mixtures is made using size exclusion, or enterokinase affinity chromatography. This step allows for desalting, buffer exchange and/or polishing, as needed.

Cleavage of the transgene product (ovalbumin-TAG-desired protein) by enterokinase, then, results in two products: ovalbumin-TAG and the desired protein.

More specific methods for isolation using the TAG label is provided in the Examples. Some desired proteins may require additions or modifications of the above-described approach as known to one of ordinary skill in the art. The method is scaleable from the laboratory bench to pilot and production facility largely because the techniques applied are well documented in each of these settings.

- In another representative embodiment, egg whites containing a protein of interest were pooled and separated, in any order, from the yolks and other egg constituents by methods known to one skilled in the art. A variety of such methods is described in manuals known in the art, such as *Egg Science & Technology*, W. Stadelman, *et al.* (Eds.), Haworth Press, Binghamton, NY (1995).

- One non-limiting example of a method for isolating a desired peptide, polypeptide or protein from an egg white is as follows. It is to be understood that this method may be employed to isolate any desired peptide, polypeptide or protein from the eggs of transgenic animals of the present invention. This present example involved transgenes that used a portion of or the entire ovalbumin protein, or specific ovalbumin epitopes, as a carrier, linked to the protein of interest via the specified TAG sequence, or another affinity/cleavage sequence. The TAG sequence contains the hairpin loop epitope from HIV I followed by an enterokinase cleavage site.

First, the viscosity of the egg white was lowered by subjecting the egg white to low shear forces of 3140 cps (Tung et al., 1969). The resulting pourable solution

- was then filtered to remove chalazae. An ammonium sulfate precipitation was then used to enrich the fraction of transgenic protein (see, for example, *Practical Protein Chemistry A Handbook* A. Darbre (Ed.), John Wiley & Sons Ltd., 1986). Other methods of crude fractionation known in the art are also used as needed.
- 5 supernatant of this separation was then fractionated using size-exclusion chromatography, further enriching the transgenic fusion protein fraction and eliminating the ammonium sulfate from the material. The fusion protein was isolated by anti-ovalbumin affinity chromatography (batch or column) using methods known to one skilled in the art. This step may capture native ovalbumin in addition to an
- 10 ovalbumin-transgene fusion protein. After elution from the anti-ovalbumin affinity resin, the transgenic protein was specifically isolated using anti-gp41 affinity chromatography (batch or column) using methods known to one skilled in the art.

Cleavage of the transgene product from the carrier and the TAG sequences was accomplished in one of at least two ways:

- 15 1) The transgenic ovalbumin-TAG-transgene target protein was left attached to the gp41 affinity resin and the protease enterokinase was added. Cleavage of the transgene by enterokinase liberated the transgene target protein from the gp41 affinity resin while the ovalbumin-TAG sequence was retained. Separation by centrifugation (in a batch process) or flow through (in a column purification), kept the transgene
- 20 target protein together with enterokinase in solution. Enterokinase was recovered and reused.

- 2) Alternatively, enterokinase was immobilized on resin (beads) by the addition of poly-lysine moieties to a non-catalytic area of the protease. The transgenic ovalbumin-TAG-transgene target protein was eluted from the gp41 affinity
- 25 chromatography resin and then applied to the protease resin. Protease action cleaved the ovalbumin-TAG sequence from the transgene target protein and left both entities in solution. The immobilized enterokinase resin was recharged and reused. The choice between these alternatives is made on a case-by case basis, depending upon the size and chemical composition of the transgene target protein.

- 30 A final separation of either of these two (process 1 or 2) protein mixtures was made using size exclusion chromatography, or enterokinase affinity chromatography. This step also allows for desalting, concentrating, buffer exchange and/or polishing, as needed.

It is believed that a typical chicken egg produced by a transgenic animal of the present invention will contain at least 0.001 mg, from about 0.001 to 1.0 mg, or from about 0.001 to 100.0 mg of exogenous protein, peptide or polypeptide, in addition to the normal constituents of egg white (or possibly replacing a small fraction of the latter). In some embodiments, a chicken egg will contain between 50 and 75 mg of exogenous protein.

One of skill in the art will recognize that after biological expression or purification, the desired proteins, fragments thereof and peptides may possess a conformation substantially different than the native conformations of the proteins, fragments thereof and peptides. In this case, it is often necessary to denature and reduce protein and then to cause the protein to re-fold into the preferred conformation. Methods of reducing and denaturing proteins and inducing re-folding are well known to those of skill in the art.

Production of Protein or Peptide in Milk

In addition to methods of producing eggs containing transgenic proteins or peptides, the present invention encompasses methods for the production of milk containing transgenic proteins or peptides. These methods include the administration of a transposon-based vector described above to a mammal through the duct system. In one embodiment, the transposon-based vector contains a transposase operably-linked to a constitutive promoter and a gene of interest operably-linked to mammary specific promoter. Genes of interest can include, but are not limited to antiviral and antibacterial proteins and immunoglobulins. In other embodiments, a transposon-based vector is administered to the ovary of an animal and germline transformation is obtained. In these embodiments, offspring of the transfected animal express a gene of interest in the mammary gland under the control of a mammary gland-specific promoter.

The following examples will serve to further illustrate the present invention without, at the same time, however, constituting any limitation thereof. On the contrary, it is to be clearly understood that resort may be had to various embodiments, modifications and equivalents thereof which, after reading the description herein, may suggest themselves to those skilled in the art without departing from the spirit of the invention.

EXAMPLE 1

IntraOviduct Administration of Transposon-Based Vectors

Quail or chicken were selected for administration of the transposon-based vectors of the present invention. Feathers were removed from the area where surgery was performed and the area was cleansed and sterilized by rinsing it with ethanol (alcohol) and 0.5% chlorhexidine. Using the scalpel, a dorsolateral incision was made through the skin over the ovary approximately 2 cm in length. Using blunt scissors, a second incision was made through the muscle between the last two ribs to expose the oviduct beneath. A small animal retractor was used to spread the last two ribs, exposing the oviduct beneath. The oviduct was further exposed using retractors to pull the intestines to one side.

A delivery solution containing a transposon-based vector and SUPERFECT® was prepared fresh immediately before surgery. Specific ratios of vector and SUPERFECT® that were used in each experiment are provided in the Examples below. The delivery solution was warmed to room temperature prior to injection into the bird. Approximately 250-500 µl of the delivery solution was injected into the lumen of the magnum of the oviduct using a 1 cc syringe with a 27 gauge needle attached. The wound was closed and antibiotic cream liberally applied to the area surrounding the wound.

20

EXAMPLE 2

Preparation of Transposon-Based Vector pTnMod

A vector was designed for inserting a desired coding sequence into the genome of eukaryotic cells, given below as SEQ ID NO:3. The vector of SEQ ID NO:3, termed pTnMod, was constructed and its sequence verified.

This vector employed a cytomegalovirus (CMV) promoter. A modified Kozak sequence (ACCATG) (SEQ ID NO:1) was added to the promoter. The nucleotide in the wobble position in nucleotide triplet codons encoding the first 10 amino acids of transposase was changed to an adenine (A) or thymine (T), which did not alter the amino acid encoded by this codon. Two stop codons were added and a synthetic polyA was used to provide a strong termination sequence. This vector uses a promoter designed to be active soon after entering the cell (without any induction) to increase the likelihood of stable integration. The additional stop codons and synthetic

30

polyA insures proper termination without read through to potential genes downstream.

- The first step in constructing this vector was to modify the transposase to have the desired changes. Modifications to the transposase were accomplished with the
- 5 primers High Efficiency forward primer (Hef) Altered transposase (ATS)-Hef 5' ATCTCGAGACCATGTGTGAACTTGATATTTTACATGATTCTCTTTACC 3' (SEQ ID NO:36) and Altered transposase- High efficiency reverse primer (Her) 5' GATTGATCATTATCATAATTCCCCAAAGCGTAACC 3' (SEQ ID NO:37, a reverse complement primer). In the 5' forward primer ATS-Hef, the sequence
- 10 CTCGAG (SEQ ID NO:38) is the recognition site for the restriction enzyme Xho I, which permits directional cloning of the amplified gene. The sequence ACCATG (SEQ ID NO:1) contains the Kozak sequence and start codon for the transposase and the underlined bases represent changes in the wobble position to an A or T of codons for the first 10 amino acids (without changing the amino acid coded by the codon).
- 15 Primer ATS-Her (SEQ ID NO:37) contains an additional stop codon TAA in addition to native stop codon TGA and adds a Bcl I restriction site, TGATCA (SEQ ID NO:39), to allow directional cloning. These primers were used in a PCR reaction with pTnLac (p defines plasmid, tn defines transposon, and lac defines the beta fragment of the lactose gene, which contains a multiple cloning site) as the template
- 20 for the transposase and a FailSafe™ PCR System (which includes enzyme, buffers, dNTP's, MgCl₂ and PCR Enhancer; Epicentre Technologies, Madison, WI). Amplified PCR product was electrophoresed on a 1% agarose gel, stained with ethidium bromide, and visualized on an ultraviolet transilluminator. A band corresponding to the expected size was excised from the gel and purified from the
- 25 agarose using a Zymo Clean Gel Recovery Kit (Zymo Research, Orange, CA). Purified DNA was digested with restriction enzymes Xho I (5') and Bcl I (3') (New England Biolabs, Beverly, MA) according to the manufacturer's protocol. Digested DNA was purified from restriction enzymes using a Zymo DNA Clean and Concentrator kit (Zymo Research).
- 30 Plasmid gWhiz (Gene Therapy Systems, San Diego, CA) was digested with restriction enzymes Sal I and BamH I (New England Biolabs), which are compatible with Xho I and Bcl I, but destroy the restriction sites. Digested gWhiz was separated on an agarose gel, the desired band excised and purified as described above. Cutting

the vector in this manner facilitated directional cloning of the modified transposase (mATS) between the CMV promoter and synthetic polyA.

To insert the mATS between the CMV promoter and synthetic polyA in gWhiz, a Stratagene T4 Ligase Kit (Stratagene, Inc. La Jolla, CA) was used and the
5 ligation set up according to the manufacturer's protocol. Ligated product was transformed into *E. coli* Top10 competent cells (Invitrogen Life Technologies, Carlsbad, CA) using chemical transformation according to Invitrogen's protocol. Transformed bacteria were incubated in 1 ml of SOC (GIBCO BRL, CAT# 15544-042) medium for 1 hour at 37° C before being spread to LB (Luria-Bertani media
10 (broth or agar)) plates supplemented with 100 µg/ml ampicillin (LB/amp plates). These plates were incubated overnight at 37° C and resulting colonies picked to LB/amp broth for overnight growth at 37° C. Plasmid DNA was isolated using a modified alkaline lysis protocol (Sambrook et al., 1989), electrophoresed on a 1% agarose gel, and visualized on a U.V. transilluminator after ethidium bromide
15 staining. Colonies producing a plasmid of the expected size (approximately 6.4 kbp) were cultured in at least 250 ml of LB/amp broth and plasmid DNA harvested using a Qiagen Maxi-Prep Kit (column purification) according to the manufacturer's protocol (Qiagen, Inc., Chatsworth, CA). Column purified DNA was used as template for sequencing to verify the changes made in the transposase were the desired changes
20 and no further changes or mutations occurred due to PCR amplification. For sequencing, Perkin-Elmer's Big Dye Sequencing Kit was used. All samples were sent to the Gene Probes and Expression Laboratory (LSU School of Veterinary Medicine) for sequencing on a Perkin-Elmer Model 377 Automated Sequencer.

Once a clone was identified that contained the desired mATS in the correct
25 orientation, primers CMVf-NgoM IV (5' TTGCCGGCATCAGATTGGCTAT (SEQ ID NO:40); underlined bases denote a NgoM IV recognition site) and Syn-polyA-BstE II (5' AGAGGTCACCGGTCAATTCTTCAGCACTGGTA (SEQ ID NO:41); underlined bases denote a BstE II recognition site) were used to PCR amplify the entire CMV promoter, mATS, and synthetic polyA for cloning upstream of the
30 transposon in pTnLac. The PCR was conducted with FailSafe™ as described above, purified using the Zymo Clean and Concentrator kit, the ends digested with NgoM IV and BstE II (New England Biolabs), purified with the Zymo kit again and cloned upstream of the transposon in pTnLac as described below.

Plasmid pTnLac was digested with NgoM IV and BstE II to remove the ptac promoter and transposase and the fragments separated on an agarose gel. The band corresponding to the vector and transposon was excised, purified from the agarose, and dephosphorylated with calf intestinal alkaline phosphatase (New England Biolabs) to prevent self-annealing. The enzyme was removed from the vector using a Zymo DNA Clean and Concentrator-5. The purified vector and CMVp/mATS/polyA were ligated together using a Stratagene T4 Ligase Kit and transformed into *E. coli* as described above.

Colonies resulting from this transformation were screened (mini-preps) as describe above and clones that were the correct size were verified by DNA sequence analysis as described above. The vector was given the name pTnMod (SEQ ID NO:3) and includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 133 -1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems), corresponding to bp 229-1873 of pGWiz. The CMV promoter was modified by the addition of an ACC sequence upstream of ATG.

Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons for stability of the transposase mRNA and for the expression of protein. More specifically, in each of the codons for the first ten amino acids of the transposase, G or C was changed to A or T when such a substitution would not alter the amino acid that was encoded.

Base pairs 2988-2993 are two engineered stop codons.

Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of 10 pGWiz.

Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from pNK2859.

Base pairs 3762 - 3831 are the 70 bp of the left insertion sequence recognized by the transposon Tn10.

Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

- 5 Base pairs 3838 - 4527 are the multiple cloning site from pBluescriptII sk(20), corresponding to bp 924-235 of pBluescriptII sk(-). This multiple cloning site may be used to insert any coding sequence of interest into the vector.

Base pairs 4528-4532 are a residue from ligation of restriction enzyme sites used in constructing the vector.

- 10 Base pairs 4533 - 4602 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 4603 - 4644 are non-coding λ DNA that is residual from pNK2859.

Base pairs 4645 - 5488 are non-coding DNA that is residual from pNK2859.

- 15 Base pairs 5489 - 7689 are from the pBluescriptII sk(-) base vector - (Stratagene, Inc.), corresponding to bp 761-2961 of pBluescriptII sk(-).

Completing pTnMod is a pBlueScript backbone that contains a *colE1* origin of replication and an antibiotic resistance marker (ampicillin).

- It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences
20 in the above construct represent restriction site remnants.

- All plasmid DNA was isolated by standard procedures. Briefly, *Escherichia coli* containing the plasmid was grown in 500 mL aliquots of LB broth (supplemented with an appropriate antibiotic) at 37°C overnight with shaking. Plasmid DNA was recovered from the bacteria using a Qiagen Maxi-Prep kit (Qiagen, Inc., Chatsworth,
25 CA) according to the manufacturer's protocol. Plasmid DNA was resuspended in 500 μ L of PCR-grade water and stored at -20°C until used.

EXAMPLE 3

Transposon-Based Vector pTnMCS

- 30 Another transposon-based vector was designed for inserting a desired coding sequence into the genome of eukaryotic cells. This vector was termed pTnMCS and its constituents are provided below. The sequence of the pTnMCS vector is provided in SEQ ID NO:2. The pTnMCS vector contains an avian optimized polyA sequence

- operably-linked to the transposase gene. The avian optimized polyA sequence contains approximately 40 nucleotides that precede the A nucleotide string.
- Bp 1 – 130 Remainder of F1 (-) ori of pBluescriptII sk(-) (Stratagene) bp1-130
- Bp 133 – 1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems) bp 229-1873
- Bp 1783 – 2991 Transposase, from Tn10 (GenBank accession #J01829) bp 108-1316
- Bp 2992 – 3344 Non coding DNA from vector pNK2859
- Bp 3345 – 3387 Lambda DNA from pNK2859
- Bp 3388 – 3457 70 bp of IS10 left from Tn10
- 10 Bp 3464 – 3670 Multiple cloning site from pBluescriptII sk(-), thru the XmaI site bp 924-718
- Bp 3671 – 3715 Multiple cloning site from pBluescriptII sk(-), from the XmaI site thru the XhoI site. These base pairs are usually lost when cloning into pTnMCS bp 717-673
- 15 Bp 3716 – 4153 Multiple cloning site from pBluescriptII sk(-), from the XhoI site bp 672-235
- Bp 4159 - 4228 70 bp of IS10 right from Tn10
- Bp 4229 - 4270 Lambda DNA from pNK2859
- Bp 4271 – 5114 Non-coding DNA from pNK2859
- 20 Bp 5115 – 7315 pBluescript sk (-) base vector (Stratagene, Inc.) bp 761-2961.

EXAMPLE 4

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/ProIns/PA) – Chicken

- 25 A vector was designed to insert a human proinsulin coding sequence under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:42.

- Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-)
- 30 (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 – 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 – 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are two engineered stop codons.

Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector
5 pNK2859.

Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-)
10 corresponding to base pairs 924-718 of pBluescriptII sk(-).

Base pairs 4050 - 4951 are a chicken ovalbumin promoter (including SDRE) that corresponds to base pairs 431-1332 of the chicken ovalbumin promoter in GenBank Accession Number J00895 M24999.

Base pairs 4958 - 6115 are a chicken ovalbumin signal sequence and
15 ovalbumin gene that correspond to base pairs 66-1223 of GenBank Accession Number V00383.1. (The STOP codon being omitted).

Base pairs 6122 - 6271 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

Base pairs 6272 – 6531 are a proinsulin gene.

Base pairs 6539 – 6891 are a synthetic polyadenylation sequence from pGWiz
20 (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6897 - 7329 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

Base pairs 7335- 7404 are the 70 base pairs of the right insertion sequence
25 (IS10) recognized by the transposon Tn10.

Base pairs 7405 - 7446 are λ DNA that is residual from pNK2859.

Base pairs 7447 – 8311 are non coding DNA that is residual from pNK2859.

Base pairs 8312 - 10512 are pBlueScript sk(-) base vector (Stratagene, Inc.)
corresponding to base pairs 761-2961 of pBluescriptII sk(-).

30

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

EXAMPLE 5

Transposon-Based Vector pTnMOD (CMV-CHOVg-ent-ProInsulin-synPA)

- A vector was designed to insert a proinsulin coding sequence under the control of a quail ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:43.

- Bp 1 – 4045 from vector pTnMod, bp 1 - 4045
- Bp 4051 – 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230-1864
- 10 Bp 5702 –6855 Chicken ovalbumin gene taken from GenBank accession # V00383, bp 66-1219
- Bp 6862 - 7011 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 7012 – 7272 Human Proinsulin taken from GenBank accession # NM000207, bp
- 15 117-377
- Bp 7273 – 7317 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)
- Bp 7318 - 7670 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271
- 20 Bp 7672 –11271 from cloning vector pTnMCS, bp 3716-7315

EXAMPLE 6

Transfection of Japanese Quail using a Transposon-based Vector containing a Proinsulin Gene via Oviduct Injections

- 25 Two experiments were conducted in Japanese quail using transposon-based vectors containing either Oval promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:42) or CMV promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:43).

- In the first experiment, the Oval promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A containing construct was injected into the lumen of the oviduct of sexually mature quail; three hens received 5 μ g at a 1:3 SUPERFECT® ratio and three received 10 μ g at a 1:3 SUPERFECT® ratio. As of the writing of the present application, at least one bird that received above-mentioned construct was producing human proinsulin in egg white (other birds remain to be tested). This

experiment indicates that 1) the DNA has been stable for at least 3 months; 2) protein levels are comparable to those observed with a constitutive promoter such as the CMV promoter; and 3) sexually mature birds can be injected and results obtained without the need for cell culture. It is estimated that each quail egg contains
5 approximately 1.4 µg/ml of the proinsulin protein. It is also estimated that each transgenic chicken egg contains 50-75 mg of protein encoded by the gene of interest.

In the second experiment, the transposon-based vector containing CMV promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A was injected into the lumen of the oviduct of sexually immature Japanese quail. A total of 9 birds were
10 injected. Of the 8 survivors, 3 produced human proinsulin in the white of their eggs for over 6 weeks. An ELISA assay described in detail below was developed to detect GP41 in the fusion peptide (Oval gene/GP41 Enterokinase TAG/Proinsulin) since the GP41 peptide sequence is unique and not found as part of normal egg white protein. In all ELISA assays, the same birds produced positive results and all controls worked
15 as expected.

ELISA Procedure: Individual egg white samples were diluted in sodium carbonate buffer, pH 9.6, and added to individual wells of 96 well microtiter ELISA plates at a total volume of 0.1 ml. These plates were then allowed to coat overnight at 4°C. Prior to ELISA development, the plates were allowed warm to room
20 temperature. Upon decanting the coating solutions and blotting away any excess, non-specific binding of antibodies was blocked by adding a solution of phosphate buffered saline (PBS), 1% (w/v) BSA, and 0.05% (v/v) Tween 20 and allowing it to incubate with shaking for a minimum of 45 minutes. This blocking solution was subsequently decanted and replaced with a solution of the primary antibody (Goat
25 Anti-GP41 TAG) diluted in fresh PBS/BSA/Tween 20. After a two hour period of incubation with the primary antibody, each plate was washed with a solution of PBS and 0.05% Tween 20 in an automated plate washer to remove unbound antibody. Next, the secondary antibody, Rabbit anti-Goat Alkaline Phosphatase-conjugated, was diluted in PBS/BSA/Tween 20 and allowed to incubate 1 hour. The plates were then
30 subjected to a second wash with PBS/Tween 20. Antigen was detected using a solution of *p*-Nitrophenyl Phosphate in Diethanolamine Substrate Buffer for Alkaline Phosphatase and measuring the absorbance at 30 minutes and 1 hour.

Additionally, a proinsulin fusion protein produced using a construct described above was isolated from egg white using ammonium sulfate precipitation and ion exchange chromatography. A pooled fraction of the isolated fusion protein was run on an SDS-PAGE gel shown in Figure 5, lanes 4 and 6. Lanes 1 and 10 of the gel contain molecular weight standards, lanes 2 and 8 contain non-transgenic chicken egg white, whereas lanes 3, 5, 7 and 9 are blank.

EXAMPLE 7

Isolation of Human Proinsulin Using Anti-TAG Column Chromatography

10 A HiTrap NHS-activated 1 mL column (Amersham) was charged with a 30 amino acid peptide that contained the gp-41 epitope containing gp-41's native disulfide bond that stabilizes the formation of the gp-41 hairpin loop. The 30 amino acid gp41 peptide is provided as SEQ ID NO:32. Approximately 10 mg of the peptide was dissolved in coupling buffer (0.2 M NaHCO₃, 0.5 M NaCl, pH 8.3 and
15 the ligand was circulated on the column for 2 hours at room temperature at 0.5 mL/minute. Excess active groups were then deactivated using 6 column volumes of 0.5 M ethanolamine, 0.5 M NaCl, pH 8.3 and the column was washed alternately with 6 column volumes of acetate buffer (0.1 M acetate, 0.5 M NaCl, pH 4.0) and ethanolamine (above). The column was neutralized using 1 X PBS. The column was
20 then washed with buffers to be used in affinity purification: 75 mM Tris, pH 8.0 and elution buffer, 100 mM glycine-HCl, 0.5 M NaCl, pH 2.7. Finally, the column was equilibrated in 75 mM Tris buffer, pH 8.0.

Antibodies to gp-41 were raised in goats by inoculation with the gp-41 peptide described above. More specifically, goats were inoculated, given a booster injection
25 of the gp-41 peptide and blood samples were obtained by veinupuncture. Serum was harvested by centrifugation. Approximately 30 mL of goat serum was filtered to 0.45 uM and passed over a TAG column at a rate of 0.5 mL/min. The column was washed with 75 mM Tris, pH 8.0 until absorbance at 280 nm reached a baseline. Three column volumes (3 mL) of elution buffer (100 mM glycine, 0.5 M NaCl, pH 2.7) was
30 applied, followed by 75 mM Tris buffer, pH 8.0, all at a rate of 0.5 mL/min. One milliliter fractions were collected. Fractions were collected into 200 uL 1 M Tris, pH 9.0 to neutralize acidic fractions as rapidly as possible. A large peak eluted from the column, coincident with the application the elution buffer. Fractions were pooled. Analysis by SDS-PAGE showed a high molecular weight species that separated into

two fragments under reducing condition, in keeping with the heavy and light chain structure of IgG.

Pooled antibody fractions were used to charge two 1 mL HiTrap NHS-activated columns, attached in series. Coupling was carried out in the same manner as that used for charging the TAG column.

Isolation of Ovalbumin-TAG-Proinsulin from Egg White

Egg white from quail and chickens treated by intra-oviduct injection of the CMV-ovalbumin-TAG-proinsulin construct were pooled. Viscosity was lowered by subjecting the allantoid fluid to successively finer pore sizes using negative pressure filtration, finishing with a 0.22 μ M pore size. Through the process, egg white was diluted approximately 1:16. The clarified sample was loaded on the Anti-TAG column and eluted in the same manner as described for the purification of the anti-TAG antibodies. A peak of absorbance at 280 nm, coincident with the application of the elution buffer, indicated that protein had been specifically eluted from the Anti-TAG column. Fractions containing the eluted peak were pooled for analysis.

The pooled fractions from the Anti-TAG affinity column were characterized by SDS-PAGE and western blot analysis. SDS-PAGE of the pooled fractions revealed a 60 kDal molecular weight band not present in control egg white fluid, consistent with the predicted molecular weight of the transgenic protein. Although some contaminating bands were observed, the 60 kDal species was greatly enriched compared to the other proteins. An aliquot of the pooled fractions was cleaved overnight at room temperature with the protease, enterokinase. SDS-PAGE analysis of the cleavage product, revealed a band not present in the uncut material that comigrated with a commercial human proinsulin positive control. Western blot analysis showed specific binding to the 60 kDal species under non-reducing condition (which preserved the hairpin epitope of gp-41 by retaining the disulfide bond). Western analysis of the low molecular weight species that appeared upon cleavage with an anti-human proinsulin antibody, conclusively identified the cleaved fragment as human proinsulin.

EXAMPLE 8

Purification Procedures for Insulin

I. ELISA data for egg characterization/identification

An ELISA was employed for the initial screening of eggs and, thereby, identification of hens producing positive eggs. With further modifications this procedure was used for the initial quantification of recombinant protein amounts. These procedures were aided by the successful purification of an initial stock of the recombinant proinsulin (RPI). This stock of protein is used in the development of a double antibody assay that increases the sensitivity and reduces the background in the assay. Subsequent identification of hens producing positive eggs obviate the need to screen each egg collected. Only periodic checks are needed to determine if production levels are consistent.

II. Egg White (EW) or Albumin Preparation

A. Clarification – Ovomucin precipitation

Eggs from hens positively identified as producing RPI are pooled for RPI purification. The initial purification step involved diluting the pool 1:1 with 100 mM Tris-HCl, pH 8 for a final concentration of 50 mM Tris-HCl. The pH of this solution was then adjusted to 6 and ovomucin was allowed to precipitate at 4°C for a minimum of 3hrs (preferably overnight) with constant stirring. The precipitated ovomucin was then pelleted and removed by centrifugation at 2400 x g. After collection of the RPI containing supernatant, the pH of this solution was readjusted to 8.

B. Filtration

To prepare the egg white for loading onto the column and, thereby, minimize the potential for clogging the columns during loading, the egg white solution was filtered to at least 0.45 um.

Initially, the ovomucin precipitated egg white solution was subjected to successive filtration steps with the pore size of the filtration membrane decreasing at each step. This procedure involved time and dilution of the egg white solution to reach 0.45 um filtration.

Amersham's hollow-fiber ultrafiltration apparatus was used to produced a column-ready solution filtered down to < 0.2 um with an undiluted starting solution. This approach minimized the time and the solution dilution needed to prepare the egg white solution for column loading.

III. Purification

A. Affinity Chromatography –

Using antibody with specificity to a synthetic peptide modeled after the enterokinase recognition site, initial purification schemes involved developing a one-step column purification procedure for the RPI.

Goats immunized with the synthetic Ent peptide were employed to produce anti-Ent Tag antiserum which was used in the egg screening ELISAs followed by antibody purification. The purified goat Anti-Ent Tag antibodies were covalently

bound to the matrix of HiTrap NHS-activated HP columns (Amersham) and subsequently used to specifically bind and purify the RPI.

5 An initial attempt was made to direct the first purification step against the ovalbumin portion of the recombinant protein using an antibody specific for the ovalbumin portion. The present purification scheme employed a combination of classical techniques such as ammonium sulfate precipitation, ion exchange, and gel filtration chromatography.

After the initial ovomucin precipitation, the egg white solution was subjected to protein precipitation using a 40% ammonium sulfate fractionation. The precipitated protein was subsequently collected via centrifugation and resuspended in 10 50 mM Tris-HCl, pH 8. The resuspended protein solution was dialyzed to remove residual $(\text{NH}_4)_2\text{SO}_4$ or subjected to gel filtration to remove the $(\text{NH}_4)_2\text{SO}_4$ and partially isolate the RPI from the remaining egg white protein. The RPI was further isolated via anion exchange chromatography using a 0 to 0.5M NaCl gradient in 50 15 mM Tris-HCl, pH 8. Two possible elution profiles were observed. One at approximately 25% of the 0.5 M NaCl gradient without $(\text{NH}_4)_2\text{SO}_4$ precipitation. The second was observed at less than 16% gradient (approximately 7%) following 40% $(\text{NH}_4)_2\text{SO}_4$ precipitation and a longer gradient. Fractions containing RPI were identified by SDS-PAGE analysis and pooled.

20 Three gel filtration columns, differing by column size and fractionation range, were employed in RPI purification and/or desalting: Superdex 75 10/300 GL, Hiload 26/60 Superdex 75, and Hiload 26/60 Superdex 200. Using these individual columns at different steps in the purification scheme increased the efficiency of the process. Fractions containing RPI were identified by SDS-PAGE analysis and pooled.

25 Cleavage of the RPI Enterokinase recognition site was accomplished using purified enterokinase from Sigma. Enterokinase, 0.004 Unit/ μl per reaction, was applied to the pooled and, if necessary, concentrated protein solution. The digestion reaction was incubated at room temperature (up to 30°C in a rolling hybridization oven) for a minimum of 16 h and in some cases up to 48 hrs of incubation. The digestion efficiency was followed using 16.5% Tris-Tricine SDS-PAGE peptide gels. 30 All gel staining utilized Simply Blue Coomassie Staining Solutions. Free Proinsulin was observed on gels after digestion.

A subsequent gel filtration separation was employed to obtain purified Proinsulin, and to remove the remaining Ovalbumin portion of the RPI and residual 35 native EW proteins. Select steps in the purification process were analyzed using the 2-dimensional Beckman Coulter ProteomeLab PF2D Protein Fractionation System.

EXAMPLE 9

Optimization of Intra-oviduct and Intra-ovarian Arterial Injections

Overall transfection rates of oviduct cells in a flock of chicken or quail hens are enhanced by synchronizing the development of the oviduct and ovary within the flock. When the development of the oviducts and ovaries are uniform across a group of hens and when the stage of oviduct and ovarian development can be determined or predicted, timing of injections is optimized to transfect the greatest number of cells. Accordingly, oviduct development is synchronized as described below to ensure that a large and uniform proportion of oviduct secretory cells are transfected with the gene of interest.

Hens are treated with estradiol to stimulate oviduct maturation as described in Oka and Schimke (T. Oka and RT Schimke, J. Cell Biol., 41, 816 (1969)), Palmiter, Christensen and Schimke (J Biol. Chem. 245(4):833-845, 1970). Specifically, repeated daily injections of 1 mg estradiol benzoate are performed sometime before the onset of sexual maturation, a period ranging from 1 – 14 weeks of age. After a stimulation period sufficient to maximize development of the oviduct, hormone treatment is withdrawn thereby causing regression in oviduct secretory cell size but not cell number. At an optimum time after hormone withdrawal, the lumens of the oviducts of treated hens are injected with the transposon-based vector. Hens are subjected to additional estrogen stimulation after an optimized time during which the transposon-based vector is taken up into oviduct secretory cells. Re-stimulation by estrogen activates transposon expression, causing the integration of the gene of interest into the host genome. Estrogen stimulation is then withdrawn and hens continue normal sexual development. If a developmentally regulated promoter such as the ovalbumin promoter is used, expression of the transposon-based vector initiates in the oviduct at the time of sexual maturation. Intra-ovarian artery injection during this window allows for high and uniform transfection efficiencies of ovarian follicles to produce germ-line transfections and possibly oviduct expression.

Other means are also used to synchronize the development, or regression, of the oviduct and ovary to allow high and uniform transfection efficiencies. Alterations of lighting and/or feed regimens, for example, cause hens to 'molt' during which time the oviduct and ovary regress. Molting is used to synchronize hens for transfection, and may be used in conjunction with other hormonal methods to control regression and/or development of the oviduct and ovary.

EXAMPLE 10

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/ProIns/PA) - Quail

- 5 A vector is designed for inserting a proinsulin gene under the control of a quail ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:44.

Base pairs 1 -130 are a remainder of F1(-) ori of pBluescriptII sk(-)
 10 (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 - 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 - 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

15 Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 - 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 - 3718 are non coding DNA that is residual from vector pNK2859.

20 Base pairs 3719 - 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 - 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 - 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

25 Base pairs 4050 - 4938 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332). Some deletions were noted in the quail sequence, as compared to the chicken sequence.

30 Base pairs 4945 - 6092 are a quail ovalbumin signal sequence and ovalbumin gene that corresponds to base pairs 54 - 1201 of GenBank accession number X53964.1. (The STOP codon being omitted).

Base pairs 6093 - 6246 are a TAG sequence containing a gp41 hairpin loop from HIV I an enterokinase cleavage site and a spacer (synthetic).

Base pairs 6247 – 6507 are a proinsulin gene.

Base pairs 6514 – 6866 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6867 - 7303 are a multiple cloning site from pBlueScriptII sk(-)
5 corresponding to base pairs 667-235 of pBluescriptII sk(-).

Base pairs 7304- 7379 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7380 - 7421 are λ DNA that is residual from pNK2859.

Base pairs 7422 – 8286 are non coding DNA that is residual from pNK2859.

10 Base pairs 8287 - 10487 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences
15 in the above construct represent restriction site remnants.

EXAMPLE 11

Preparation of Transposon-Based Vector pThMod(Oval/ENT TAG/p146/PA) – Chicken

20 A vector was designed for inserting a p146 gene under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird. The vector sequence is provided below as SEQ ID NO:45.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

25 Base pairs 133 – 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 – 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

30 Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector pNK2859.

Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

- 5 Base pairs 4050 - 4951 are a chicken ovalbumin promoter (including SDRE, steroid-dependent response element) that corresponds to base pairs 431-1332 of the chicken ovalbumin promoter in GenBank Accession Number J00895 M24999.

- Base pairs 4958 - 6115 are a chicken ovalbumin signal sequence and Ovalbumin gene that correspond to base pairs 66-1223 of GenBank Accession
10 Number V00383.1 (The STOP codon being omitted).

Base pairs 6122 - 6271 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

Base pairs 6272 – 6316 are a p146 sequence (synthetic) with 2 added stop codons.

- 15 Base pairs 6324 – 6676 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6682 - 7114 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

- Base pairs 7120- 7189 are the 70 base pairs of the right insertion sequence
20 (IS10) recognized by the transposon Tn10.

Base pairs 7190 - 7231 are λ DNA that is residual from pNK2859.

Base pairs 7232 – 8096 are non coding DNA that is residual from pNK2859.

Base pairs 8097 - 10297 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

25

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

30

EXAMPLE 12

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/p146/PA) – Quail

A vector was designed for inserting a p146 gene under the control of a quail ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird. The vector sequence is given below as SEQ ID NO:46.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 - 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

5 Base pairs 1780 - 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 - 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922-2337 of pGWiz.

10 Base pairs 3415 - 3718 are non coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 - 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

15 Base pairs 3838 - 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

Base pairs 4050 - 4938 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332).

20 Bp 4945 - 6092 are a quail ovalbumin signal sequence and ovalbumin gene that corresponds to base pairs 54 - 1201 of GenBank accession number X53964.1. (The STOP codon being omitted).

Base pairs 6097 - 6246 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

25 Base pairs 6247 - 6291 are a p146 sequence (synthetic) with 2 added stop codons.

Base pairs 6299 - 6651 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

30 Base pairs 6657 - 7089 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

Base pairs 7095- 7164 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7165 - 7206 are λ DNA that is residual from pNK2859.

Base pairs 7207 – 8071 are non coding DNA that is residual from pNK2859.

Base pairs 8072 - 10272 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

- 5 It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

EXAMPLE 13

10 *Additional Transposon-Based Vectors for Administration to an Animal*

- The following example provides a description of various transposon-based vectors of the present invention and several constructs that have been made for insertion into the transposon-based vectors of the present invention. These examples are not meant to be limiting in any way. The constructs for insertion into a transposon-based vector are provided in a cloning vector pTnMCS or pTnMod, both described above.

pTnMCS (CMV-CHOVg-ent-ProInsulin-synPA) (SEQ ID NO:47)

- Bp 1 – 3670 from vector PTnMCS, bp 1 - 3670
- 20 Bp 3676 – 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864
- Bp 5327 – 6480 Chicken ovalbumin gene taken from GenBank accession # V00383, bp 66-1219
- Bp 6487 – 6636 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- 25 Bp 6637 – 6897 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377
- Bp 6898 – 6942 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)
- 30 Bp 6943 - 7295 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271
- Bp 7296 – 10895 from cloning vector pTnMCS, bp 3716-7315

pTnMCS (CMV-prepro-ent-ProInsulin-synPA)

Bp 1 – 3670 from vector pTnMCS, bp 1 - 3670

Bp 3676 – 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864

- 5 Bp 5326 - 5496 Capsite/prepro taken from GenBank accession # X07404, bp 563–733
- Bp 5504 - 5652 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 5653 – 5913 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377

- 10 Bp 5914 – 5958 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)
- Bp 5959-6310 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271
- Bp 6313–9912 from cloning vector pTnMCS, bp 3716-7315

15

pTnMCS(Chicken OVep+OVg'+ENT+proins+syn polyA)

Bp 1–3670 from vector pTnMCS, bp 1 - 3670

Bp 3676–4350 Chicken Ovalbumin enhancer taken from GenBank accession #S82527.1 bp 1-675

- 20 Bp 4357–5692 Chicken Ovalbumin promoter taken from GenBank accession # J00895M24999 bp 1-1336
- Bp 5699–6917 Chicken Ovalbumin gene from GenBank Accession # V00383.1 bp 2-1220. (This sequence includes the 5'UTR, containing putative cap site, bp 5699-5762.)
- 25 Bp 6924–7073 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 7074–7334 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 7335-7379 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- 30 Bp 7380-7731 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 7733–11332 from vector pTnMCS, bp 3716 - 7315

pTnMCS(Chicken OVep+prepro+ENT+proins+syn polyA)

- Bp 1 – 3670 from cloning vector pTnMCS, bp 1 - 3670
- Bp 3676 – 4350 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675
- 5 Bp 4357 – 5692 Chicken Ovalbumin promoter taken from GenBank accession # J00895-M24999 bp 1-1336
- Bp 5699–5869 Cecropin cap site and prepro, Genbank accession # X07404 bp 563-733
- Bp 5876 - 6025 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- 10 Bp 6026 - 6286 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 6287 - 6331 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- Bp 6332 - 6683 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- 15 Bp 6685 – 10284 from cloning vector pTnMCS, bp 3716 - 7315

pTnMCS(Quail OVep+OVg'+ENT+proins+syn polyA)

- Bp 1 – 3670 from cloning vector pTnMCS, bp 1 - 3670
- 20 Bp 3676 – 4333 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession # S82527.1, bp 1-675. (There are multiple base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)
- 25 Bp 4340 – 5705 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)
- 30 Bp 5712 – 6910 Quail Ovalbumin gene, EMBL accession # X53964, bp 1-1199. (This sequence includes the 5'UTR, containing putative cap site bp 5712-5764.)
- Bp 6917 - 7066 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 7067 - 7327 Human proinsulin GenBank Accession # NM000207 bp 117-377

- Bp 7328 - 7372 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- Bp 7373 - 7724 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- 5 Bp 7726 - 11325 from cloning vector pTnMCS, bp 3716 - 7315
- pTnMCS(Quail OVep+prepro+ENT+proins+syn polyA)
- Bp 1 - 3670 from cloning vector pTnMCS, bp 1 - 3670
- Bp 3676 - 4333 Quail Ovalbumin enhancer: 658 bp sequence, amplified from quail
- 10 genomic DNA, roughly equivalent to the far- upstream chicken ovalbumin enhancer, GenBank accession #S82527.1, bp 1-675. (There are multiple base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)
- Bp 4340 - 5705 Quail Ovalbumin promoter: 1366 bp sequence, amplified from quail
- 15 genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)
- Bp 5712-5882 Cecropin cap site and prepro, Genbank accession # X07404 bp 563-
- 20 733
- Bp 5889 - 6038 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 6039 - 6299 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 6300 - 6344 Spacer DNA, derived as an artifact from the cloning vectors pTOPO
- 25 Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- Bp 6345 - 6696 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 6698 - 10297 from cloning vector pTnMCS, bp 3716 - 7315.

- 30 pTnMOD (CMV-prepro-ent-proins-synPA)
- Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045
- Bp 4051 - 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230-1864
- Bp 5701-5871 Capsite/prepro taken from GenBank accession # X07404, bp 563-733

- Bp 5879 - 6027 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 6028-6288 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377
- 5 Bp 6289 - 6333 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)
- Bp 6334 - 6685 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271
- Bp 6687 -10286 from cloning vector pTnMCS, bp 3716-7315
- 10 pTnMOD(Chicken OVep+OVg'+ENT+proins+syn polyA)
- Bp 1 - 4045 from cloning vector pTnMod, bp 1 - 4045
- Bp 4051 - 4725 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675
- 15 Bp 4732 - 6067 Chicken Ovalbumin promoter taken from GenBank accession # J00895-M24999 bp 1-1336
- Bp 6074 - 7292 Chicken Ovalbumin gene from GenBank Accession # V00383.1 bp 2-1220. (This sequence includes the 5'UTR, containing putative cap site bp 6074-6137.)
- 20 Bp 7299 - 7448 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 7449 - 7709 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 7710 - 7754 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- 25 Bp 7755 - 8106 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 8108 - 11707 from cloning vector pTnMod, bp 3716 - 7315
- pTnMOD(Chicken OVep+prepro+ENT+proins+syn polyA)
- 30 Bp 1 - 4045 from cloning vector pTnMCS, bp 1 - 4045
- Bp 4051 - 4725 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675
- Bp 4732 - 6067 Chicken Ovalbumin promoter taken from GenBank accession # J00895-M24999 bp 1-1336

- Bp 6074–6244 Cecropin cap site and prepro, Genbank accession # X07404 bp 563-733
- Bp 6251 - 6400 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- 5 Bp 6401 - 6661 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 6662 - 6706 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- Bp 6707 - 7058 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- 10 Bp 7060 – 10659 from cloning vector pTnMCS, bp 3716 - 7315
- pTnMOD(Quail OVep+OVg⁺+ENT+proins+syn polyA)
- Bp 1 – 4045 from cloning vector pTnMCS, bp 1 - 4045
- Bp 4051 – 4708 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house
- 15 from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession # S82527.1, bp 1-675. (There are multiple base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)
- Bp 4715 – 6080 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house
- 20 from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)
- Bp 6087 – 7285 Quail Ovalbumin gene, EMBL accession # X53964, bp 1-1199. (This
- 25 sequence includes the 5'UTR, containing putative cap site bp 6087-6139.)
- Bp 7292 - 7441 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 7442 - 7702 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 7703 - 7747 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- 30 Bp 7748 - 8099 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 8101 – 11700 from cloning vector pTnMCS, bp 3716 - 7315

pTnMOD(Quail OVEp+prepro+ENT+proins+syn polyA)

Bp 1 – 4045 from cloning vector pTnMCS, bp 1 - 4045

Bp 4051 – 4708 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession #S82527.1, bp 1-675. (There are multiple base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)

Bp 4715 – 6080 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)

Bp 6087–6257 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-733

Bp 6264 - 6413 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6414 - 6674 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 6675 - 6719 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 6720 - 7071 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 7073 – 10672 from cloning vector pTnMCS, bp 3716 - 7315

pTnMOD (CMV-prepro-ent-hGH-CPA)

Bp 1–4045 from vector PTnMOD, bp 1 - 4045

Bp 4051–5694 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230-1873

Bp 5701-5871 Capsite/Prepro taken from GenBank accession # X07404, bp 563–733

Bp 5878-6012 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6013–6666 Human growth hormone taken from GenBank accession # V00519, bp 1-654

Bp 6673–7080 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058

Bp 7082–10681 from cloning vector pTnMOD, bp 4091–7690

pTnMCS (CHOVep-prepro-ent-hGH-CPA)

Bp 1–3670 from vector PTnMCS, bp 1–3670

- 5 Bp 3676–4350 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1, bp 1–675
Bp 4357–5692 Chicken Ovalbumin promoter taken from GenBank accession # J00899-M24999, bp 1–1336
Bp 5699–5869 Capsite/Prepro taken from GenBank accession # X07404, bp 563–733
- 10 Bp 5876–6010 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
Bp 6011–6664 Human growth hormone taken from GenBank accession # V00519, bp 1–654
Bp 6671–7078 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651–11058
- 15 Bp 7080–10679 from cloning vector pTnMCS, bp 3716–7315

pTnMCS (CMV-prepro-ent-hGH-CPA)

Bp 1 – 3670 from vector PTnMCS, bp 1 - 3670

- 20 Bp 3676–5319 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230–1873
Bp 5326–5496 Capsite/Prepro taken from GenBank accession # X07404, bp 563 – 733
Bp 5503–5637 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- 25 Bp 5638–6291 Human growth hormone taken from GenBank accession # V00519, bp 1–654
Bp 6298–6705 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651–11058
Bp 6707–10306 from cloning vector pTnMCS, bp 3716–7315

30

pTnMOD (CHOVep-prepro-ent-hGH-CPA)

Bp 1–4045 from vector PTnMOD, bp 1–4045

Bp 4051–4725 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1, bp 1–675

- Bp 4732-6067 Chicken Ovalbumin promoter taken from GenBank accession # J00899-M24999, bp 1-1336
- Bp 6074-6244 Capsite/Prepro taken from GenBank accession # X07404, bp 563-733
- Bp 6251-6385 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added
 5 enterokinase cleavage site
- Bp 6386-7039 Human growth hormone taken from GenBank accession # V00519, bp 1-654
- Bp 7046-7453 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058
- 10 Bp 7455-11054 from cloning vector pTnMOD, bp 4091-7690
- PtMod(CMV/Transposase/ChickOvep/prepro/ProteinA/ConpolyA)
- BP 1-130 remainder of F1 (-) ori of pBluescriptII sk(-) (Stratagene) bp 1-130.
- BP 133-1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy
 15 Systems) bp 229-1873.
- BP 1780-2987 Transposase, modified from Tn10 (GenBank #J01829).
- BP 2988-2993 Engineered DOUBLE stop codon.
- BP 2994-3343 non coding DNA from vector pNK2859.
- BP 3344-3386 Lambda DNA from pNK2859.
- 20 BP 3387-3456 70bp of IS10 left from Tn10.
- BP 3457-3674 multiple cloning site from pBluescriptII sk(-) bp 924-707.
- BP 3675-5691 Chicken Ovalbumin enhancer plus promoter from a Topo Clone 10 maxi 040303 (5' XmaI, 3' BamHI)
- BP 5698-5865 prepro with Cap site amplified from cecropin of pMON200 GenBank #
 25 X07404 (5'BamHI, 3'KpnI)
- BP 5872-7338 Protein A gene from GenBank# J01786, mature peptide bp 292-1755 (5'KpnI, 3'SacII)
- BP 7345-7752 ConPolyA from Chicken conalbumin polyA from GenBank # Y00407 bp 10651-11058. (5'SacII, 3'XhoI)
- 30 BP 7753-8195 multiple cloning site from pBluescriptII sk(-) bp 677-235.
- BP 8196-8265 70 bp of IS10 left from Tn10.
- BP 8266-8307 Lamda DNA from pNK2859
- BP 8308-9151 noncoding DNA from pNK2859
- BP 9152-11352 pBluescriptII sk(-) base vector (Stratagene, INC.) bp 761-2961

All patents, publications and abstracts cited above are incorporated herein by reference in their entirety. It should be understood that the foregoing relates only to preferred embodiments of the present invention and that numerous modifications or
5 alterations may be made therein without departing from the spirit and the scope of the present invention as defined in the following claims.

CLAIMS

We Claim:

- 5 1. A method of producing a transgenic animal comprising, administering to an oviduct or an ovary of an animal a composition comprising a transposon-based vector.
- 10 2. The method of claim 1, wherein the composition is injected into an artery leading to the oviduct or the ovary.
3. The method of claim 1, wherein the composition is injected into a lumen of the oviduct.
- 15 4. The method of claim 1, wherein the composition further comprises a transfection reagent.
5. The method of claim 1, wherein the transposon-based vector comprises:
 - 20 a) a transposase gene operably linked to a first promoter, the transposase gene encoding for a transposase; and
 - b) one or more genes of interest operably-linked to one or more additional promoters; wherein the one or more genes of interest and their operably-linked promoters are flanked by transposase insertion sequences recognized by the transposase, and wherein the first promoter comprises a
25 modified Kozak sequence comprising ACCATG (SEQ ID NO:1).
- 30 6. The method of claim 5, wherein one to twenty codons at a beginning of the transposase gene are modified by changing a nucleotide at a third base position of the codon to an adenine or thymine without modifying the amino acid encoded by the codon.
7. The method of claim 1, wherein the transposon-based vector comprises:
 - 35 a) a transposase gene operably linked to a first promoter and an avian optimized polyA sequence, the transposase gene encoding for a transposase; and
 - b) one or more genes of interest operably-linked to one or more additional promoters;

c) wherein the one or more genes of interest and their operably-linked promoters are flanked by transposase insertion sequences recognized by the transposase.

5 8. The method of claim 7, wherein the first promoter is a constitutive promoter.

9. The method of claim 7, wherein the first promoter is an oviduct-specific promoter selected from the group consisting of ovalbumin, ovotransferrin, ovomucoid, ovomucin, g2 ovoglobulin, g3 ovoglobulin, 10 ovoflavoprotein, and ovostatin.

10. The method of claim 7, wherein the one or more gene of interest is operably-linked to a second promoter.

11. The method of claim 10, wherein the second promoter is an oviduct-specific promoter selected from the group consisting of ovalbumin, ovotransferrin, ovomucoid, ovomucin, g2 ovoglobulin, g3 ovoglobulin, 15 ovoflavoprotein, and ovostatin.

12. The method of claim 7, wherein the transposon-based vector further comprises an egg directing sequence or an enhancer operably-linked to the one or more genes of interest.

13. A method of increasing expression of a polynucleotide sequence in an avian comprising, administering to the avian a composition comprising the polynucleotide sequence operably-linked to an avian optimized polyA sequence.

14. A transgenic animal made with the method of Claim 1.

15. The transgenic animal of Claim 14, wherein the animal is a bird.

16. An egg produced by the transgenic bird of Claim 15.

17. The egg of Claim 16, wherein the egg contains a protein, a polypeptide or a peptide encoded by the transposon-based vector.

18. A method of producing proteins, polypeptides or peptides comprising:
- a) administering to an oviduct of an egg-laying animal a composition comprising a transposon-based vector, wherein transposon-based vector comprises a transposase gene operably linked to a first promoter, the transposase gene encoding for a transposase, and one or more genes of interest operably-linked to one or more additional promoters; and,
 - b) permitting the one or more genes of interest to be expressed into a protein, a polypeptide or a peptide.
19. The method of Claim 18, further comprising
- a) collecting an egg from the egg-laying animal;
 - b) harvesting egg white containing the protein, the polypeptide or the peptide; and,
 - c) purifying the protein, the polypeptide or the peptide.
20. The method of any of the preceding claims, wherein the animal is a bird.
21. The method of any of the preceding claims, wherein the animal is a poultry bird.
22. The method of any of the preceding claims, wherein the transposase is a Tn10 transposase.
23. A method of producing a transgenic mammal comprising, administering to an intramammary duct system of a mammal a composition comprising a transposon-based vector.
24. A transgenic mammal made with the method of Claim 23.
25. Milk produced by the transgenic mammal of Claim 24.
26. The milk of Claim 25, wherein the milk contains a protein, a polypeptide or a peptide encoded by the transposon-based vector.
27. A method of producing proteins, polypeptides or peptides comprising:
- a) administering to an intramammary duct system of a mammal a composition comprising a transposon-based vector, wherein transposon-based

vector comprises a transposase gene operably linked to a first promoter, the transposase gene encoding for a transposase, and one or more genes of interest operably-linked to one or more additional promoters; and,

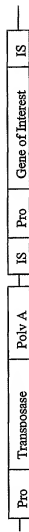
- 5 b) permitting the one or more genes of interest to be expressed into a protein, a polypeptide or a peptide.

28. The method of Claim 27, further comprising

- 10 a) collecting milk from the mammal; and,
 b) purifying the protein, the polypeptide or the peptide.

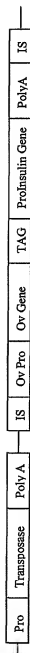
1/5

FIGURE 1



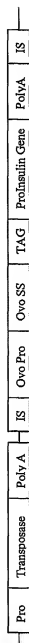
2/5

FIGURE 2



3/5

FIGURE 3



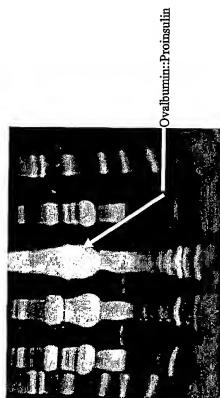
4/5

FIGURE 4

IS	Tet, Pro	Ovgen	Pro	Ovotrans	Pro	Ovomucin	IS
----	----------	-------	-----	----------	-----	----------	----

5/5

FIGURE 5



Appendix A

SEQ ID NO:1 (modified Kozak sequence)
 ACCATG

5

SEQ ID NO:2 (pTnMCS)

1 ctgacgcgcc ctgtagcgcc gcatlaagcg cggcggggtgt ggtgggttac cgcagcgtga
 61 ccgcgtacaact tgcacgcgcc ctagcgccgc ctcccttgcg tttcttccct tccctttctcg
 121 ccaagcttgcg cggcatcaga ttggctattg gccattgcat aogtgtgac catatataaa
 181 tatgtacatt tatattggct catgtccaac attacgcgca tggtagacat gattattgac
 241 tagttattaa tagttaatcaa ttacgggggtc attagttaac agcccatata tggagttcgc
 301 cgttacaata cttacggtaa atggccgcgc tggctgacgc cccaagcacc ccgcgccatt
 361 gacgtcaata atgacgtatg ttcccatagt aacgcgaata gggactttcc atgacgcgta
 421 atcggtgggag tatttacggg aaacttgcca cttggcagta catcaaggtc atcatatgcc
 481 aagtaacgcc cttattgacg tcaattgacg taaattggccc gctcgggaatt atgcocagta
 541 catgaactta tgggaacttcc ctacttgcca gtacatctac gtattaggta tgcctattac
 601 catggttgatg cgggtttggc agtaactcaa tggcgctgga tagcggtttg acctacgggg
 661 attccaagt ctccacccca ttgaagtcga tgggagtttg ttctggcacc aaatacaagc
 721 ggaactttcca aaatgtcgta acacctccgc cccattgacc caaatggggcg gttagcggtgt
 781 aeggtgggag gtcatataaa gcagagctcg tttagtgacg cgtcagatcg cctggagagc
 841 ccaatcagac tattttgacc tccacagaag aacccggagc cgtaccagcc tccgcgggcg
 901 ggaaagctgac attggaacgc ggaattcccg tgcacagagt gacgtaaagta cgcctatag
 961 actcctatagg cacaccocct tggctcttat gcatgctata cgttttttgg cttggggcct
 1021 ataacacccc gottccttat gctatagggt atggtatagc ttacgctata ggtgtggggt
 1081 attgacccatt attgaacact cccctattgg tgaacatact ttccattact aaatcataac
 1141 atggtctctt gccacaacta tctctattgg ctatatgcca atactctgc ctcacagac
 1201 tgcacgcgcat tctgtatttt tacaggatgg ggtccattct attattaca aaatcataac
 1261 tcaacaacgc cgtcccccgc tgcgcgcagt ttttatataa catagcgtgg agcttccagc
 1321 cgaactctcgg gtagctgttc cggacatggg cttctctcgc gttagcgccg gatcccaac
 1381 tccagagccct ggtcccatgc ctccagcgcc tcatggtcgc tcgggagctc ctgtctccta
 1441 acagtggagg ccaagcttag gcacagcaca atgcaccaca cccagctggt gcgcgacaga
 1501 gccgctggcg taggttatgt gtctgaaaaa gagcgtggag attgggctcg cacgcgtcac
 1561 gcagatggaa gacttaagcg agcggcagaa gaagatgcag gcagctgagt tgttgtattc
 1621 tgataagagt cagaggtaac tcccggttgc gtgctgttaa cggctggagg cagtgtatgc
 1681 tgcagcagta tctgtgtcgc cgcgcgcgcc accagacata atagctgaga cactgaacga
 1741 ctgttctctt ccatgggtct ttctgcagt caccgtcgga ccatgtgga actcgtattt
 1801 ttaacagact ctctttacca attctgcgcc gaattacact taaaacgact caacagctta
 1861 aogttggctt gccgcgcat acttgactgt aaaaactcca ctcttacga atttggcgtg
 1921 aaocctgcga ccaagcgag aacaaaaaat aacatacaac gaatgcgacc atgttgagc
 1981 aatcgtcacc tccacaaga gcgactcgct gtataccggt ggcactgtag ctcttatctg
 2041 tcgggcaata cgtatggcct tgtacttgtt gactgtctgt atattctgta gccaaacga
 2101 cttatgggat tgcgagcttc agtgcgacta cagcgtcgtt ctgttactct tatatgaaa
 2161 gcttctcgcg ttccagagca atgttcaaa agagctcatg accaatttct agccgacctt
 2221 gcgagcattc taacagagta caccacacgc ctacattgtc gtgactgtcg ctcttaagct
 2281 cactgggtata aatccgttga gaagctgggt tggtagctgt taagtgcagg aagaggaaaa
 2341 gtacaatatt cagccttagg agcggaaaac tggaaaacct taagcaactt acatgatgtg
 2401 tcatctagtc actcaagac tttagcttat aagaggtcga ctaaaagcaa tccaactcca
 2461 tgcocaaatt tattgtataa atctcgtctc aaagcgcgaa aaaatcagcg ctgcagacgt
 2521 acctactgtc accaccgcgc actcaaaatc tactcagcgt cggcgaagga gccactgggt
 2581 ctagcaacta acttactcgt tgaat tcca acaccccaac aactgtttaa tctctatctg
 2641 aagcgaatgc agattgaaga aacottccga gaacttcaaa gtctcgctca cgtactaggc
 2701 ctacgcgata gccgaacagc cagctcagag cgttttgata tcatgtctgt aatcgccctg
 2761 atgcttcaac taacatggtg gcttgcgggc gtctcagctg agaaacagg ttgggaacaag
 2821 caactccagc ctacacagc cagaatctga aacgtactct caacagtctg cttaggcctg
 2881 gacttttgc ggaattctgc ctacacata acaggggaag acttactcgt ggcctgaacc
 2941 ctactagctc aaattttat ctacacatgg tactcgtttg ggaattatg aggggatcgc
 3001 cttagagcga tccgggatct cgggaaaagc gttgggtgacc aaagtgccct ttgggaacaag
 3061 ctttaaaat aaaaaacaa tactcatgct cgtttataag cagcaatata ttatgattga
 3121 tgcctacatc acaaaaaaaa ctgatttaac ctgatttaac taatataaaa cttatccta
 3181 tgaacatatt ctgtattata ttattgtata atgaacttga aaaaaattag ctttgaatac
 3241 gtagcctata attggttga atgttatgac aagagaacaa acttaagact tctcagactg
 3301 aggttaaacgc catgttcagc aaattgatcc aagagaacaa acttaagact tctcagactg
 3361 aatggttaatt ctggttgacc ctgagcactg atgaatcccc taatgatttt ggttaaaaac
 3421 attaagttaa ggtgtgata catctgtcca tatgatcccg gtaattgtg tatgcctact
 3481 cattaaggcac ccaagctttt acactttatg acactttatg tcatgtgtgt tgaattgtg

3541 agcggataac aatttcacac aggaaaacagc tatgacacatg attacgccaa gogcgcaatt
 3601 aacccctcact aaagggaacaa aaagctggag ctccacgcgog gtggcgggcg ctctagaact
 3661 agtggatccc ccgggctgcga ggaattcgat atcaacgctta tgcataccgc tgacctcgag
 5 3721 gggggggccg gtaccocaatt cgcctatag tggatcgtat taacgogcgt cactggcgct
 3781 cgtttttacaa cgtcgtgact gggaaaaacc tggcgtttacc caacttaatc gcttcgtcag
 3841 acatccccct ttccgcagct gcgctaagat cgaagagccc cgcaccgatc gcccttccca
 3901 acagtctgcg agcctgaatg gogaattggaa atttgaagcg ttaatatatt ctgtaaaattc
 3961 cggttaaat ttgtttaaat cagctcaatt tttaaccaat aggcgcaaat cggcaaaatc
 4021 ccttataaat caaagaata gacogagata ggggttgagt ttgttcocgt ttggacaagc
 10 4081 agtcacctat taagaacaagt ggaactccaa gctcaaggcg gaaaaacgtc ctctaggcgt
 4141 gatggccacc tatctccggga tcatatgaca agatgtgat ccaccttaac ttaattgatt
 4201 ttaaccaaat cattagggga ttcatcagtg ctcaagggtca acgagaatta acattcogct
 4261 aggaagactt atgatgatga tgtgcttaaa accttactca atggcgtggt atgcattatcg
 4321 caatacatgc gaaaaaccta aaagagcttg ccgataaaaa aggcacattt attgctattt
 15 4381 acgcgogctt tttattgagc ttgaagaata aataaaatag ataggttttt ttggaagcta
 4441 aatctctctt atcgttaaaaa atgcctcctt ggggttatcaa gagggctcat atattctcgc
 4501 gaataacatc atttggtgac gaaatacctt agcacttgct tctgtgttac tccctgagc
 4561 ttgagggggtt aacatgaagg tcatcgatag caggataata atacagttac tccctaaocg
 20 4621 aataactcaa accacgcact cccaatttgg tagtgaatga agggctcaac tccctaaocg
 4681 taatgggcca ataacacggg ttgcatttgg aggtctcaac tctctgttgc tccctaaocg
 4741 ttctcgatga ctctttgtttt ggaatagacat cactccctgt aatgcaggtg aggcagaccc
 4801 accacacgac aataaatta ttgcgaataa tccgaataaa actgcogttca taaacgctaa
 4861 aaaggaaaaat gcactactat ctgcgaataa tccgaacagt actgcogttt attgtcccat
 25 4921 ttatggctca tcttctcctgc cacaaggctt tggaaactgt agtgtaaaag accaagaccc
 4981 gtaatgcaaaa gccaaacatc atgctattca tcatcagat ttctgtataa gccacaccc
 5041 gctcgtgatt ggctatcaat cgcctgaagt aataatcaac aaatggcatc gttaaatag
 5101 tgaattatac cgcactgact ttgttccctt tagtgaaggt taatgcgog cttagggtaa
 5161 tcatgctgaat agctgtttcc ttgttgaaat ttgtatccgc tcaaatcc acacacata
 30 5221 cgaagccgaa gcaataaagt taaagcctgg ggtgcctaat gagtgaacta actcaatta
 5281 attcogtctgc gctcaactgc cgtttcccg tcgggaaacc tctgtgcga cgtcactaa
 5341 tgaactgcgc aaacgcgggg gagagcggtt ttgcgtattt ggogctcttc cgtctccatg
 5401 ctcaactgact cgtcgcgttc ggtcgttcgg ctgcgcgag ccgtatacgc tcaatcaaa
 5461 ggcgttaata ggttatccac agaactcagg gataacgcag gaaagaacat ctgagcaaaa
 35 5521 ggcgcacaaa aggcacaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataagctc
 5581 cgcctccctg acagatca caaaaactca gctcgaagtc agaggtggcg aaacccgcga
 5641 ggaactataaa gataccagcc gtttccctct ggaagctccc tctgtgcctc tctctgtcc
 5701 accctgcgcg ttaccggata cctgcgcgc ttctcctct cgggaagctt ggcgtctctg
 5761 catagctcac gctgtaggta tctcagttcg gtgtaggctt ttctgcctaa gctggcgctg
 40 5821 gtgcacgaac ccccggttca gccgcacgcg tgcgcctat tgcgtaacta ccggtcttag
 5881 tccaaacccg taagacaaga cttatgcoca ctggcagcag ccaactgtaa caggattagc
 5941 agagcgaggt atgtaggcgg tgcctacagag ttcttgaagt ggttgctcaa caggtgtac
 6001 actgaaggga cagtatttgg tatctgcgct ctgctgaagc cagttacott ccggaaaaaa
 45 6061 gttggtgact ctgtatccg caacaaaacc acgcgtggt cgggtgtgtt ttttgttgc
 6121 aagcagcaga ttacgcgcag aaaaaaggga tctcaaggag atcctttgat cttttctcag
 6181 gggctctgag ctcaagtga cgaaaaactca cgttaaaggga ttttggtctg gagattatca
 6241 aaaggagatc tcaactagat ccttttaaat taaaaa tga gttttaact aatctaaagt
 6301 atatatgagt aaacttggtc tgaacagttac caatgcttaa tcaagtgggc acctactca
 6361 ggcgtctgct tatttctgtc atccatagtt gctgaactcc ccgtcgtgta gataactcag
 40 6421 aacggggagg gcttaacatc tggcccccag gctgcaatga tacocgcaga cccocgtca
 6481 ccgctgcocag atttatcagc aataaaacag ccagccggaa gggccgagcg cagaagctgt
 50 6541 cgtcgaactt tatccgcctc catccagctt attaatgtt gctgcaatga ggcgggtgag
 6601 agctgcocag ttaaatggtt ggcgcacggt gttgccattg ctacaggcat cgtgtgtga
 6661 cgtctcoggt ttggtatggc tcatctcagc ttcatctcag ctoogtctcc atcogtatca cgtgtgtga
 55 6721 tgaatccoca tggttgtgaa aaaaagcggt agctccttgc gctcctcag atctgttact
 6781 agtaagttgg ccgcagtggt atcactcatg gttatggcag cactgcataa tttcttact
 6841 gtcactgcat ccgttaagat cttttctgtg actgtgagt actcaocaa gtcacttga
 6901 gaatagtgta tggcggaacc gagttgctct tgcccgcggt caatacggga taatacog
 6961 ccaactagca gaactttaaa agtgcctcct attgaaaaac gttctctggg gogaaaaact
 7021 taaagtagct tacogctgtt gagatccagt tgcgtgtaac ccaactcgtc aggcgaana
 60 7081 tcttcagcat ctttactctt caccagcggt tctgggtgag caaaaacagc aagcgcaaat
 7141 ccgcgcaaaa agggaataag gggcgacagg aaatgttga tactcatact ctctctttt
 7201 caatatattt gaagcattta tcagggttat tgcctatgga ggggatacat atttgaagt
 7261 atttagaaaa ataaacaat aggggttcog gcacatttc ccgaaaaagt gccac

SEQ ID NO:3 (pTnMod)

	CTGACGCGCC	CTGTAGCGGC	GCATTAAAGCG	CGCGGGGTGT	GGTGGTTACG	50
	CGCAGCGTGA	CCGCTACACT	TGCCAGCGCC	CTAGCGCCCG	CTCCTTTCGC	100
	TTTCTTCCCT	TCCTTTCTCG	CCACGTTTCGC	CGGCATCAGA	TTGGCTATTG	150
	GCCATTGCAT	ACGTTGTATC	CAIATCATAA	TATGTACATT	TATATTGGCT	200
5	CATGTCCAAC	ATTACCGCCA	TGTTGACATT	GATTATTGAC	TAGTATTATA	250
	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT	AGCCCATATA	TGGAGTTCGC	300
	CGTTACATAA	CTTACGTTAA	ATGGCCCGCC	TGGCTGACCG	CCCAAGGACC	350
	CCCGCCCAAT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AAOCCCAATT	400
	GGGACTTTCC	ATTGACGCTA	ATGGGTGGAG	TATTTACCGT	AAACTGCCCC	450
10	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTACGCCC	CCTATTGACG	500
	TCAATGACGG	TAAATGGCCC	GCCTGGCATT	ATGCCCACTA	CATGACCTTA	550
	TGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TGCTATTATC	600
	CATGTGTGATG	CGGTTTTTGG	AGTACATCAA	TGGGCGTGGA	TAGCGGTTTG	650
	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TGACGCTCAA	TGGGAGTTTG	700
15	TTTTGGCACC	AAAATCAACG	GGACTTTTCCA	AAATGTGCTA	ACAACCTCCG	750
	CCCATGTGACG	CAAATGGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	800
	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCAAGC	850
	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGCCG	900
	GGAACTGGTC	ATTGGAAACG	GGATTCCCCG	TGCCAAGAGT	GACGTAAATA	950
20	CCGCTCTATG	ACTCTATAGG	CACACCCCTT	TGCGTCTTAT	GCAIGCTATA	1000
	CTGTTTTTTG	CTTGGGGCCT	ATACACCCCC	GCTTCCTTAT	GCTATTAGGT	1050
	ATGGTATAGC	TTAGCCTATA	GGTGTGGGTT	ATTGACCATT	ATTGACCATT	1100
	CCCTTATTGG	TGACGATACT	TTCCATTACT	ATCCCATAAC	ATGGCTCTTT	1150
	GCCACAACCTA	TCTCTATTGG	CTATATGCCA	ATACTCTGTC	CTTCAGAGAC	1200
25	TGACACGGAG	TCGTATTATT	TACAGGATGG	GGTCCCATTT	ATTATTATTA	1250
	AATTCACATA	TACAACAACG	CGCTCCCCCG	TCCGCGCAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAATCTCGG	GTACGTGTTC	CGGACATGGG	1350
	CTCTTCTCGG	GTAGCGGGCG	AGCITCCACA	TCCGAGCCCT	GGTCCCATCT	1400
	CTCCAGCGGC	TCATGTGCGC	TCGGCAGCTC	CTTGCTCCTA	ACAGTGGAGG	1450
30	CCAGACTTAG	GCACAGCACA	ATGCCCAACA	CCACCAGTGT	CCGCGCAACG	1500
	GCCTGTGGCG	TAGGGTATGT	GTCTGAAAT	GAGCGTGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGGAA	GACTTAAAGC	AGCGGCAGAA	GAGATGCGAG	1600
	GCAGCTGAGT	TGTTGTATT	TGATAAGAGT	CAGAGGTAA	TCCGTTGGCG	1650
	GTGCTGTTAA	CGGTGGAGGG	CAGTGTAGTC	TGAGCAGTAC	TGCTTCTCTG	1700
35	CGCGCGCGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCCTTT	1750
	CCATGGGTCT	TTTCTGCAAT	CACCGTCGGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGAIT	CTCTTTACCA	ATTCGCCCC	GAATTACACT	TAAAACGACT	1850
	CAACAGCTTA	ACGTGGCCTT	GCCAGCGATT	ACTTGACTGT	AAAACCTCTA	1900
	CTCTTACCAG	ACTTGGCCGT	AACTTGCCAA	CCAAAGCGAG	AACAAAACAT	1950
40	AACTCAAAAC	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAAGA	2000
	GGGACTCGCT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TGGGAATAC	2050
	GATGCCCAT	GTACTTGTGT	ACTGGTCTGA	TATTCTGTAG	CAAAAACGAC	2100
	TTATGGTATT	GCAGCTTCA	GTGCACTAC	ACGGTCGTT	TGTTACTCTT	2150
45	TATGAAAAG	CGTTCCCGCT	TTGAGAGCAA	TGTTCAAAGA	TAAGCTCATGA	2200
	CCAATTCTTA	GCCGACCTTG	CGAGCATTCT	ACCGAGTAAC	ACCACCCCG	2250
	TCATTGTGAG	TGATGCTGCG	TTTAAAGTGC	CATGGTATAA	ATCCGTTGAG	2300
	AAAGCTGGGT	GGTACTGGTT	AAAGTCAGTA	AGAGGAAAAG	TACAATATGC	2350
	AGACCTAGGA	GGGGAACACT	GGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
	CATCTAGTCA	CTCAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAGCAAT	2450
50	CAATCTCAT	GCCAAATCT	ATTGTATAAA	TCTGCTCTTA	AAGCGCAAAA	2500
	AAATCAGCGC	TGCACGCGGA	CTCATGTGCA	CCACCCGTCA	CCTAAAATCT	2550
	ACTCAGCGTC	GGCAAGGAG	CCATGGGTT	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTCGAA	CACCCAAACA	ACTGTTTAA	ATCTATTGCA	AGGAAATGCA	2650
	GATTGAAGAA	ACCTTCCGAG	ACTTGAAGAA	TCCTGCGCTAC	GGACTAGGCC	2700
55	TACGCCATAG	CGGAACGAGC	AGCTCAGAGC	GTTTTGATAT	CATGCTGCTA	2750
	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	CTTGGGGGCG	TTATGCTCTA	2800
	GAAACCAAGGT	TGGGACAGAC	ACTTCAGGCG	TAAACACGTC	AGAAATCGAA	2850
	ACGTACTCTC	AACAGTTGCG	TTAGGCATGG	AAGTTTTGCG	GCAATCTGCG	2900
	TACACAATAA	CAAGGGAGA	CTTACTCGTG	GCTGCAACCC	TACTAGCTCA	2950
60	AAATTATTCT	ACACATGGTT	ACGCTTTGGG	GAAATTATGA	TAATGATCCA	3000
	GATCACTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTGTGTGG	3050

	TTTTTGTGG	ATCTGCTGTG	CCTTCTAGTT	GCCAGCCATC	TGTTGTTTGC	3100
	CCCTCCCCCG	TGCTTCTCCT	GACCTTGGAA	GGTGCCACTC	CCACTGTCCCT	3150
	TTCTCTAATA	AATGAGGAAA	TGTCATCGCA	TGTGTCGAGT	AGGTGTCTAT	3200
	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACA	GCAGGGGGGA	GGATTGGGAA	3250
5	GACAATAGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAGG	TGCTGAAGAA	3400
	TTGACCCGGT	GACCAGAGGT	GCCTTTTATC	ATCACTTTAA	AAATAAAAAA	3450
	CAATTACTCA	GTGCTGTGTA	TAAGCAGCAA	TTAATTATGA	TTGATGCCTA	3500
10	CATCACAACA	AAAACCTGAT	TAACAAATGG	TGTGTCGCC	TTAGAAAGTA	3550
	TATTTGAACA	TATCTTGAT	TATATTATG	ATAATAATAA	AAACCTTATC	3600
	CCATATCCAAG	AAGTGATGCC	TATCATTTGG	TGGAATGAAC	TTGAAAAAAA	3650
	TTAGCCTTGA	ATACATTACT	GGTAAGGTAA	ACGCCATTGT	CAGCAAAATG	3700
	ATCCAGAGGA	ACCAACTTAA	AGCTTTCCCT	ACGGAATGTT	AATTCTCGTT	3750
15	GACCCGTAGC	ACTGATGAAT	CCCCTAATGA	TTTTGGTAAA	AATCATTAG	3800
	TTAAGGTGGA	TACACATCTT	GTCATATGAT	CCCGTAATG	TGAGTTAGCT	3850
	CATCACTTAG	GCACCCGAGG	CTTTACACTT	TATGCTTCGG	GCTCGTATGT	3900
	TGTGTGGAAAT	TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC	3950
20	CATGATTACG	CCAAGCGCGC	AATTAACCTT	CACATAAGGG	AACAAAGCT	4000
	GGAGCTCCAC	CGCGGTGGCG	GC CGCTCTAG	AACATAGTGA	TCCC CGCGGC	4050
	TGCAGGAATT	CGATATCAAG	CTTATCGATA	CGCGTGACCT	CGAGGGGGGG	4100
	CCCGGTACCC	AATTGGCCCT	ATAGTGAGTC	GTTATTACGG	CGCTCACTGG	4150
	CCGTGCTTTT	ACAACGTCGT	GACTGGGAAA	ACCCCTGGCGT	TCCCAACTAT	4200
	AATCGCCTTG	CAGCACATCC	CCCTTTCCGC	AGCTGGCGTA	ATAGCGAAGA	4250
25	GGCCCGCACC	GATGCGCCCT	CCCAACAGTT	CGCGAGCCTG	AATGGCGAAT	4300
	GGAAATTTGA	AGCGTTAATA	TTTTTGTTAAA	ATTGCGTTTA	AATTTTGTG	4350
	AAATCAGCTC	ATTTTTTAAC	CAATAGGCCG	AAATCGGCAA	AATCCCTTAT	4400
	AAATCAAAAG	AATGACCGCA	GATAGGGTTG	AGTGTGTTC	CAGTTTGGAA	4450
	CAAGAGTCCA	CTATTAAAGA	ACGTGGACTC	CAACGTCAA	GGGCGAAAAA	4500
30	CGCTCTATCA	GGCGATGGC	CCACTACTCC	GGGATCATAT	GACAGAAATG	4550
	GTATCCACCT	TAACTTAATG	ATTTTTTACCA	AAATCATTAG	GGGATTCATC	4600
	AGTGCTCAGG	GTCAACGAGA	ATTAACATTC	CGTCAGGAAA	GCTTATGATG	4650
	ATGATGTGCT	TAAAACTTA	CTCAATGGCT	GGTTATGCAT	ATCGCAATAC	4700
	ATTCGAAAAA	CCCTAAAGAG	CTTCGCGATA	AAAAAGGCCA	ATTTATTGCT	4750
35	ATTGTACCGG	GCTTTTATT	GAGCTTGAAA	GATAAATAAA	ATAGATAGGT	4800
	TTTATTTGAA	GCTAAATCTT	CTTTATCGTA	AAAAATGCC	TCTTGGGTTA	4850
	TCAAGAGGGT	CATTATATTT	CGCGGAATAA	CATCAATTGG	TGACGAAATA	4900
	ACTAAGCACT	TGCTCTCTGT	TTACTCCCTT	GAGCTTGAGG	GGTTAACTAT	4950
	AGGTCTATCG	ATAGCAGGAT	AATAATACAG	TAAACCGTGA	AACCAATAAT	5000
40	CCAAATCCAG	CCATCCCAAA	TTGGTAGTGA	ATGATTATAA	ATAACAGCAA	5050
	ACCAATATGG	GCCAATAACA	CCGGTTGCAAT	TGGTAAGGCT	CACCAATAAT	5100
	CCCTGTAAAG	CACCTTGCTG	ATGACTCTTT	GTTTGGATAG	ACATCACTCC	5150
	CTGTAAATGCA	GGTAAAGCGA	TCCCACCACC	AGCCAATAAA	ATTAATAACG	5200
	CGAAAACTAA	CCAACTTCA	GATATAAAGC	CTAAAAAGGC	AAATGCACTA	5250
45	CTATCTGCAA	TAAATCCGAG	CAGTACTGCC	GTTTTTTCGC	CCATTTAGTG	5300
	GCTATTCTTC	CTGCCACAAA	GGCTTGGAAAT	ACTGAGTGTA	AAAGACCAAG	5350
	ACCCGTAAATG	AAAAGCCAAC	CATCATGTCTA	TTCATCATCA	CGATTTCTGT	5400
	AATAGCACCA	CACCGTGTG	GATTGGCTAT	CAATCGCGTG	AAATAATAAT	5450
	CACAACATGG	CATCGTTAAA	TAAGTGATGT	ATACCGATCA	GCTTTTGTCT	5500
50	CCTTTAGTGA	GGGTTAATGT	CGCGTTGGC	GTAATCATGG	TCATAGCTGT	5550
	TTCTGTGTG	AAATGTGTAT	CGCTCACAA	TTCCACACAA	CATACGAGCC	5600
	GGAGCATAA	AGTGAATAAG	CTGGGGTGCC	TAAAGTGA	GCTAACTCAT	5650
	TTAATTTGGG	TTGCGCTCAC	TGCCCGCTTT	CCAGTCGGGA	AACCTGTCTG	5700
	GCCAGCTGCA	TTAATGAATC	GGCCACCGCG	CGGGGAGAGG	CGGTTTGGCT	5750
55	ATTGGGCGCT	CTTCGCTTCT	CTCGCTCACT	GACTCGCTGC	GCTCGGTGCT	5800
	TCGGCTCGCG	CGAGCGGTAT	CAGCTCACTC	AAAGCGCGTA	ATACGGTTAT	5850
	CCACAGAAATC	AGGGGATAAC	CGAGGAAAGA	ACATGTGAGC	AAAAAGCCAG	5900
	CAAAAGGCCA	GGAAACCGTAA	AAAGCCCGCG	TGTCTGGCGT	TTTTCCATAG	5950
	GCTCCGCCCC	CCTGACGAGC	ATCAAAAAAA	TCGACGCTCA	AGTCAGAGGT	6000
60	GGCGAAACCC	GACAGGACTA	TAAAGATACC	AGGCGTTTCC	CCCTGGAAGC	6050
	TCCCTCGTGC	GCTCTCTGT	TCCGACCCTG	CGCGTTACCG	GATACCTGTC	6100

CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA 6150
 GGTATCTCAG TTCGGTGTAG GTCGTTGCTT CCAAGCTGGG CTGTGTGCAC 6200
 GAACCCCCCG TTACAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT 6250
 TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG 6300
 5 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG 6350
 AAGTGGTGGC CTAACCTACG CTACACTAGA AGGACAGTAT TTGGTATCTG 6400
 CGCTCTGTCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT 6450
 CCGGCARACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTSCAAGCAG 6500
 CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC 6550
 10 TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG 6600
 TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT AATTAATAAA 6650
 TGAATTTTAA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG 6700
 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGCTATTTTC 6750
 GTTTCATCAT AGTTGCTCTG CTCGCCGTCG TGATAGATAAC TAAGATACGG 6800
 15 GAGGCGTTAC CATCTGGCCC CAGTGTCTGA ATGATACCGC GAGACCCACG 6850
 CTCACCGGCT CCAGATTTAT CAGCAATAAA CCAGCCACGC GGAAGGCGCC 6900
 AGCCGAGGAG TGGTCTCTGA ACTTTATCCG CCTCCATCCA GTCTATTAAAT 6950
 TGTTCGCCGG AAGCTAGAGT AAGTAGTTCG CCMGTAAATA GTTTGCGCAA 7000
 CGTTGTTGCC ATTGCTACAG GCATGTTGGT GTCACGCTCG TCGTTTGCTA 7050
 20 TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGCGAGT TACATGATCC 7100
 CCCATTGTTT GCAAAAAAGC GGTAGCTCC TTGCGTCTCT CGATCGTTGT 7150
 CAGAAGTAAG TTGGCCGCAG TGTTATCACT CATGTTATG GCAGCACTGC 7200
 ATATTTCTCT TACTGTCTAT CCATCCGTAA GATGCTTTTC TGTGACTGGT 7250
 GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG 7300
 25 CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAAGTT 7350
 TAAAAGTGCT CATCATTTGA AAACGTTCTT CGGGCGGAAA ACTCTCAAGG 7400
 ATCTTACCGC TGTGTAGATC CAGTTCGATG TAACCCACTC GTGCACCAAA 7450
 CTGATCTTCA GCATCTTTTA CTTTACCAG CGTTTCTGGG TGAGCAAAAA 7500
 CAGGAGGCGA AATGCCGCA AAAAAGCGAA TAAGGGGAC ACAGAAATGT 7550
 30 TGAATACTCA TACTCTTCTT TTTTCAATAT TATTGAAGCA TTTATCAGGG 7600
 TTATTGTCTC ATGACGGGAT ACATATTTAG ATGTATTTAG AAAATAAAC 7650
 AAATAGGGGT TCCGGGCACA TTTCCCGGAA AAGTGCCAC 7689

35 SEQ ID NO:4 (a Kozak sequence)
ACCATGG

SEQ ID NO:5 (a Kozak sequence)
ACCATGT

40 SEQ ID NO:6 (a Kozak sequence)
AAGATGT

SEQ ID NO:7 (a Kozak sequence)
ACGATGA

45 SEQ ID NO:8 (a Kozak sequence)
AAGATGG

50 SEQ ID NO:9 (a Kozak sequence)
GACATGA

SEQ ID NO:10 (a Kozak sequence)
ACCATGA

55 SEQ ID NO:11 (conalbumin polyA)
ctgtccattg ctgtctctc tgccttctct cgtcactctg aatgtggctt ctctgactact
gccacagcaa gaaataaaat ctcaacatct aatgtgggtt cctgagggtt ttcaagagtc

gttaagcacat ttccttcccc agcaccacct gctgcaggcc agtgccaggc accaacttgg
ctactgctgc ccatgagaga aatccagttc aatatatttc aaagcaaat ggattacata
tgccctagat cctgattaac aggcgtttgt attatctagt gctttcgctt caccagatt

§

SEQ ID NO:12 (synthetic polyA)
GGCGCTCGATCGAGATCACCTCTGCTGCTAAAGATCAGAGCTCTAGAGATCTGTGTGTTGGTTTTT
TGTGGATCTGCTGCTGCTCTAGTGTGCACGCCATCTGTGTGTTGGCCCTCCCCCGTGCCTCTGTACC
CTGGAAGAGTGCACCTACCGCATCTGCTCTTAATAAATAGAGAAATGCATCGCTGTCTAGTAGAG
TGTCTATCTTAATTCTGGGGGGTGGGTGGTGGCCAGCAGCAGAGGGGGAGATTTGGGAAGACAAATGACAGG
CATGCTGGGAGTGGGTGGGCTCTATGGGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCGGTACCTCTCTC

10

SEQ ID NO:13 (avian optimized polyA)

15

ggggatcgc tctagagcga
tccgggtatc cgggaaaaagc
ctctgggtgacc aaaggctgacct
tttatcatcaca tctctaaaaat
aaaaaacatat tactcactgct
ctgttataaagc cagcacaattaa
ttatgatgtga tgcctacatat
acacacaaaaa ctgtatttaac
aaatgctgttg tctgccttat
aaagtatatat tgacacattat
cttgattatat ttattgataa
taataaaaaac cttatcccta
tccagaagtgat gctgcctac
attggttgata atgaacttga
aaaaaaattag acttgaatac
attactcgtaa cggtaaacgc
cattgtcagc aaattgatcc
aagagaccaa a

20

aacacacgaac
 ctgttataag
 tcatatgaga
 aaacacaaata
 aaatggtttgg
 aaagtgtatt
 ctgttatata
 taataaaaaa
 tccaagaagt
 attggtttga
 aaaaaattga
 attacttgta
 cattgtcagc
 aaggaaacca
 a

25

cttggattata ttattgataa
taataaaaaac cttatcccta
tccaagaagt gatgcctatc
attgggtgga atgaacttga
aaaaaattag ccttgaatac
attactggta aggtaaacgc
cattgtcagc aaattgatcc
aaagagaacca a

30

aaaaatttag cctcgatcde
 attactggta aggtaaacgc
 cattgtcagc aaattgatcc
 aagaqaacca a

35

SEQ ID NO:14
(vitellogenin promoter)

40

TGAATGTGTT	CTGTGGTTAT
CAATATAAAT	CACAGITAGT
GATGAAGTGT	GCTGCAAGCC
TGCATCAGTT	GCTGCAAGTT
GCTGCAATTT	GTATTTTGGT
CTGTAGGAAA	TGCAAAAGGT
CTAGGCTGTA	TGCCACATTC
TATCCCTCTT	GCCTTATGCT
TGAGAATCTC	TGCAGGTTTT
AATTGTCFAC	ATTTTGGCTC
CATTTACTTT	GGAAGATATAA
AATTTCACAG	TAATCTTTATG
AAACCTTTGT	CACTTTAAAAA
ATATTCTCGT	TCAGGCTGAC
CGGAGCTGAA	ACAGACACAT
GATCCCGTGA	TTCTAATAAA
TACATATGTT	CCATATATGT
TTTCTCAGTA	GCCCTCTAAAT
CTATGTGCGT	TGGTGCACTA
ATGAATACAT	GAATTAGCAAAA
GGTTTATCTG	GATTAGCTGC
TGGCTCCGAG	GAATGGCCAT
TAACCCAAAG	TGAATGGAGA

45

50

55

60

GGGAGAGTAT AGTCAATGTA
 GATTATACTG ATTGCTGATT
 GGGTTATTAT CAGCTAGATA
 5 ACAACTGGG TCAGGTGCCA
 GGTCAACATA ACCTGGGCAA
 AACCAGTCTC ATCTGTGGCA
 GGACCATGTA CCAGCAGCCA
 GCCGTGACCC AATCTAGGAA
 AGCAAGTAGC ACATCAATTT
 10 TAAATTTATT GTAAATGCCG
 TAGTAGAAGT GTTTTACTGT
 GATACATTGA AACTTCTGGT
 CAATCAGAAA AAGGTTTTTT
 ATCAGAGATG CCAAGGTATT
 15 ATTTGATTTT CTTTATTGCG
 CGTGAAGAGA ATTTATGATT
 GCAAAAGAG GAGTGTTTAC
 ATAAACTGAT AAAAACTTG
 AGGAATTCAG CAGAAACAG
 20 CCACGTGTTT CTGAACATTC
 TTCCATAAAA GTCTCACCAT
 GCCTGGCAGA GCCTATTCA
 CCTTCGGT
 25
 SEQ ID NO:15 (fragment of ovalbumin promoter - chicken)
 GAGGTCAGAA GTTTCCTTTA CTGTTTGTC ATTCTATTAT TTCAATACAG
 AACATAGCT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG
 30 TCCTTCGAAC CATGAACACT CTCACGCTG AATTTCACAA TTCCTCTGTC
 ATCTGCCAGG CCATTAAAGT ATTCATGGAA GATCTTTGAG GAACACTGCA
 AGTTTCATATC ATAAACACAT TTGAAATGA GTATTGTTT GCATTGTATG
 GAGCTATGTT TTGCTGTATC CTCAGAAAAA AAGTTTGTTA TAAAGCATTC
 ACACCCATAA AAGATAGAT TTAATATTTC CAGCTATAGG AAAGAAAGTG
 35 CGTCTGCTCT TCACTCTAGT CTCAGTTGGC TCCTTCACAT GCATGCTTCT
 TTATTTCCTC TATTTTGTCA AGAAAATAAT AGGTCACGTC TTGTTCTCAC
 TTATGCTCTG CCTAGCATGG CTCAGATGCA CGTTGTAGAT ACAAGAAGGA
 TCAAAATGAA CAGACTTCTG GTCTGTTACT ACAACCATAG TAATAAGCAC
 ACTAACTAAT AATTGCTAAT TATGTTTCC ATCTCTAAGG TTCCACATTC
 40 TTTCTGTTTT CTAAAGATC CCATTATCTG GTTGTAAGT AAGCTCAATT
 GAACATGAGC AATATTTCCC AGTCTTCTCT CCCATCCAAC AGTCTCTGAT
 GATTAGCAGA ACAGGCGCAA AACACATTGT TACCCAGAAT TAAAACTAA
 TATTTGCTCT CCATTCAATC CAAATGGAC TATTGAAAC TAAATCTRA
 CCCAATCCCA TTAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG
 45 GGAAOCTGTG GGTGGGTGAC AATTACAGCT ATATATTCCC CAGGCGTCAG

SEQ ID NO:16 (chicken ovalbumin enhancer)
 ccgggctgca gaaaatgcc aggtggacta tgaactcaca tccaaaggag
 50 cttgacctga tacctgattt tcttcaaact ggggaaacaa cacaatccca caaaacagct
 cagagagaaa ccatcactga tggctacagc acaagggtat gcaatggcaa tccattcgac
 attcatctgt gacctgagca aaatgattta tctctccatg taatggttgc tcttccctc
 attgaaaggc aatttccaca ctccacaata gcaacaaga caaacagaga acaattatg
 tgcctcttc taatgtcaaa attgtagtgg caaaggagg acaaaaatct caagttctga
 55 gtaggtttta gtgattggat aagaggcttt gacctgtgag ctccactgga ctccatattc
 tcttggataa aaagtgcctt tataactttc aggtctccga gtcttatttc atgagactgt
 tgggttaggg acagaccac aatgaatgc ctggcatagg aaaggggcag agagccttag
 ctgacctttt ctgggacaa gcaattgcaa acaatgtgtg acaaaactat ttgtactgct
 60 ttgcacagct gtgctgggca gggcaatcca ttgccaccta tcccaggtaa ccttccaact
 gcaagaagat tgttgcctac tctctctaga

SEQ ID NO:17 (5' untranslated region)

GTGGATCAACATACAGCTAGAAAGCTGTATTGCTTTAGCACTCAAGCTCAAAAGACAAC TCAGAGTTC
ACC

5

SEQ ID NO:18 (putative cap site)

ACATACAGCTAG AARGCTGTAT TGCCITTAGC ACTCAAGCTC AAAAGACAAC TCAGAGTTCA

10

SEQ ID NO:19 (Chicken Ovalbumin Signal Sequence)

ATG GGTCCATCG GCGCAGCAAG CATGGAATTT TGTITTTGATG TATTCAAGGA GCTCAAAGTC
CACCATGCCA ATGAGAACAT CTTCTACTGC CCCATTGCCA TCATGTGAGC TCTAGCCATG
GTATACCTGG GTGCAAAAGA CAGCACCAGG ACACAGATAA ATAAGGTTGT TGGCTTTGAT
AAACTTCCAG GATTGCGAGA CAGTATTGAA GCTCAGTGTG GCACATCTGT AAAAGTTCAC

15

TCTTCCATTG GAGACATCCT CAACCAAATC ACCAAACCAA ATGATGTTTA TTGCTTCAGC
CTTGCCAGTA GACTTTATGC TGAAGAGAGA TACCCAATCC TGCCAGAATA CTGTCAGTGT
GTGAAGGAAC TGTATAGAGG AGGCTTGGAA CCTATCAACT TTCAAACAGC TGCAGATCAA
GCCAGAGAGC TCATCAATTC CTGGGTAGAA AGTCAGACAA ATGGAATTAT CAGAATAATGTC

20

CTTCAGCCAA GCTCCGTGGA TTCTCAAAC TCAATGGTTC TGGTTAATGC CATTGTCTTC
AAAGGACTGT GGGAGAAAC ATTTAAGGAT GAAGACACAC AAGCAATGCC TTTCAGAGTG
ACTGAGCAAG AAGCAAACC TGTGCAGATG ATGTACCAGA TTGGTTTATT TAGAGTGGCA
TCAATGGCTT CTGAGAAAT GAAGATCCTG GAGCTTCCAT TTGCCAGTGG GACAATGAGC

25

ACTGTTGTGC TGTTCCTGA TGAAGTCTCA GGCCTTGAGC AGCTTGAGAG TATAATCAAC
TTTGAAAAAC TGACTGAATG GACCAGTTCT AATGTTATGG AAGAGAGGAA GATCAAAATG
TACTTACTCT GCATGAAGAT GGAGGAAAAA TACAACCTCA CATCTGCTCT AATGGCTATG
GGCATTACTG ACGTGTITAG CTCTTCAGCC AATCTGTCTG GCATCTCCTC AGCAGAGAGC

30

CTGAAGATAT CTCAGCTGT CCATGCAGCA CATGCAGAAA TCAATGAAGC AGGCAGAGAG
GTGGTAGGGT CAGCAGAGGC TGGAGTGGAT GCTGCAAGCG TCTCTGAAGA ATTTAGGGCT
GACCATCAT TCCTCTTCTG TATCAAGCAC ATCGCAACCA ACGCCGTTCT CTCTTTGGC
AGATGTGTTT CCCCT

35

SEQ ID NO:20 (Chicken Ovalbumin Signal Sequence - shortened 50bp)

ATG GGTCCATCG GCGCAGCAAG CATGGAATTT TGTITTTGATG TATTCAAGGA

40

SEQ ID NO:21 (Chicken Ovalbumin Signal Sequence - shortened 100bp)

ATG GGTCCATCG GCGCAGCAAG CATGGAATTT TGTITTTGATG TATTCAAGGA GCTCAAAGTC
CACCATGCCA ATGAGAACAT CTTCTACTGC CCCATTGCCA

45

SEQ ID NO:22 (vitellogenin targeting sequence)

ATGAGGGGATCATACTGGCATTAGTGCTCACCCCTGTAGCGCAGCAAGTTTGACATTGGT

50

SEQ ID NO:23 (pro-insulin sequence)

TTTGTGAACCAACCTGTGCGGCTCACACCTGGTGGAGCTCTCTACCTAGTGTGCGGGGAACGAGGC
TCTTCTTACACACCAAGACCCTGCGGAGGAGCAGAGGACCTGCAGTGGGGCAGGTGAGCTGGGCGGG
GGCCCTTGGTGCAGGCGCTGCGAGCCCTTGGCCCTGGAGGGTCCCTGCAGAGCGTGGCATTTGGAA
CAATGCTGTACCAGCATCTGCTCCCTTACCAGCTGGAGAACTCTGCAACTAG

55

SEQ ID NO:24 (p146 protein)

KYKALKKIAKLL

60

SEQ ID NO:25 (p146 coding sequence)

AAATACAAAAAGCACTGAAAAAAGCTGGCAAAAGTCTGCT

65

SEQ ID NO:26 (spacer)

(GPGG)_x

SEQ ID NO:27 (spacer)
GPGGGPGGGPGG

5

SEQ ID NO:28 (spacer)
GGGGSGGGSGGGGS

10

SEQ ID NO:29 (spacer)
GGGGSGGGSGGGSGGGGS

15

SEQ ID NO:30 (repeat domain in TAG spacer sequence)
Pro Ala Asp Asp Ala

20

SEQ ID NO:31 (TAG spacer sequence)
Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp
Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp

25

SEQ ID NO:32 (gp41 epitope)
Ala Thr Thr Cys Ile Leu Lys Gly Ser Cys Gly Trp Ile Gly Leu Leu

30

SEQ ID NO:33 (polynucleotide sequence encoding gp41 epitope)
Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Thr Thr Cys Ile Leu Lys Gly
Ser Cys Gly Trp Ile Gly Leu Leu Asp Asp Asp Lys

35

SEQ ID NO:34 (enterokinase cleavage site)
DDDDK

40

SEQ ID NO:35 (TAG sequence)
Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp
Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Thr Thr Cys Ile Leu Lys Gly Ser Cys
Gly Trp Ile Gly Leu Leu Asp Asp Asp Lys

45

SEQ ID NO:36 (altered transposase Hef forward primer)
ATCTCGAGACCATGTGTGAACCTTGATATTTTACATGATTCTCTTTACC

50

SEQ ID NO:37 (altered transposase Her reverse primer)
GATTGATCATTATCATAATTCCCAAGCGTAACC

55

SEQ ID NO:38 (Xho I restriction site)
CTCGAG

SEQ ID NO:39 (Bcl I restriction site)
TGATCA

SEQ ID NO:40 (CMVf-NgoM IV primer)
TTGCCGGCATCAGATTGGCTAT

5

SEQ ID NO:41 (Syn-polyAr-BstE II primer)
AGAGGTACC CGGGTCAATTCTTCAGCACTGGTA

10	SEQ ID NO:42 (pTnMod(Oval/ENT tag/Proins/PA) - Chicken)	
	CTGACGCGCC CTGTAGCGGC GCATTAAAGC CGGCGGGTGT GGTGGTTACG	50
	CGCAGCGTGA CCGCTACACT TGCCAGCGCG CTAGCGCGCG CTCCTTTCCG	100
	TTCTTCCTCT TCCTTTCTCG CCAAGTCCGC CGGCATCAGA TTGGCTATTG	150
	GCCATTGTCAT ACGTGTGATC CATATCATAA TATGTACATT TATATTGGCT	200
15	CATGTCCAAAC ATTACGCCCA TGTGTGACATT GATTATTGAC TAGTTATTAA	250
	TAGTAATCAA TTACGGGGTC ATTAGTTTAT AGCCCATATA TGGAGTTCGG	300
	CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC	350
	CCGCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA	400
	GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACCGGT AAACCTGCCCA	450
20	CTTGCGAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG	500
	TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA	550
	TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTAC TCGCTATTAC	600
	CATGTGTATG CGGTTTTCGC AGTACATCAA TGGCGTGGGA TAGCGGTTTG	650
	ACTCACGGGG ATTTCCAAAT CTCCACCCCA TTGACGTCAA TGGGAGTTTG	700
25	TTTGGGACC AAAATCAACG GGACTTTCCA AAATGTGCTA ACAACTCCGC	750
	CCCATGTAGC CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA	800
	GCAGAGCTCG TTTAGTGAA CCGTCAGATCG CCTGGAGACG CCATCCACGC	850
	TGTTTTTACC TCCATAGAAG ACACCGGGAG CGATCCAGCC TCCGCGCGCG	900
	GAACCGGTGC ATTGGAACGC GGATTTCCCG TGCCAAGAGT GACGTAAGTA	950
30	CCGCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGTCTATA	1000
	CTGTTTTTGG CTGCGGGCCT ATACACCCCG GCTTCCTTAT GCTATAGGTT	1050
	ATGGTATAGC TTAGCCTATA GGTGTGGGTT ATTGACCATT ATTGACCATT	1100
	CCCTTATGG TGAAGTACT TTTCCATTACT AATCCATAAC ATGGCTCTTT	1150
	GCCACAACCTA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTCAGAGAC	1200
35	TGACACGGAC TCTGTATTTT TACAGGATGG TGGTCCATT TTTATTATTA	1250
	AATTACATA TACAACAACG CCGTCCCGCG TGCCCGCAGT TTTTATTAAA	1300
	CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTC CGGACATGGG	1350
	CTCTTCTCCG GTAGCGGGCG AGCTTCCACA CTGCGAGCCCT GGTCCCATGC	1400
	CTCCAGCGGC TCATGTGCGC TCGGCAGTCT TTGTCTCTA ACAGTGGAGG	1450
40	CCAGACTTAG GCACAGCACA ATGCCCAACA CCACAGTGT GCCGCACAAG	1500
	GCGGTGCGGG TATGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGGCTCG	1550
	CACGCTGTAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCGA	1600
	GCAGCTGAGT TGTGTATTTC TGATAAGAGT CAGAGGTAAC TCCCGTTGGG	1650
	GTGCTGTATA CGGTGAGAGG CAGTGTAGTC TGAGCAGTAC TCGTGTGCTGC	1700
45	CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTCTT	1750
	CCATGGGTCT TTTCTGCACT CACCGTGGGA CCTGTGTGA ACTTGATATT	1800
	TTACATGATT CTCTTTACCA ATTCTGCCCC GAATTACACT TAAACGACTT	1850
	CACAGCTTA ACGTGTGCTT GCCACGCACT ATTGACTGT AAAACTCTCA	1900
	CTCTTACCGA ACTTGGCGGT AACCTGCCAA CCAAGCGAG AACAAACAT	1950
50	AACATCAAA CAAATGACCG ATTGTAGGT AATCGTCACC TCCACAAAGA	2000
	GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC	2050
	GATGCCCAAT GTACTGTTG ACTGCTCTGA TATTCGTGAG CAAAAACGAC	2100
	TTATGGTATT GCGAGCTTCA GTGCGACTAC ACGGTCTGTC TGTTACTCTT	2150
	TATGAGAAAG CGTTCGCCCT TTCAGACGAA TGTTCAAAGA AAGCTCATGA	2200
55	CCAATTTCTA GCGGACCTTG CGAGCATTTCT ACCAGGTAAC ACCACACCG	2250
	TCATTGTTCAG TGAATGCTGC TTTAAAGTGC CATGGTATAA ATCCGTTGAG	2300
	AAGCTGGGTT GGTACTGGTT AAGTCGATGA AGAGGAAAG TACAATATGC	2350
	AGACTTAGGA GCGGAAACT GGAAACCTAT CAGCAACTTA CATGATATGT	2400
	CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAGCAAT	2450
60	CCAATCTCAT GCCAAATCTT ATTGTATATA TCTGCTCTA AAGGCCGAAA	2500
	AAATCAGCGC TCGACACGGA CTCATTGTCA ACCCGCTCA CCTAAATCT	2550

	ACTCAGCGTC	GGCAAAGGAG	CCATGGGTTT	TAGCAACTAA	CTTACCTGTT	2600
	GAATTCGAA	CACCCAAACA	ACTTGTAAAT	ATCTATTGCA	AGCGAATGCA	2650
	GATTGGAAGA	ACCTTCCGAG	ACTTGAAGAA	TCCTGCCTAC	GGACTAGGCC	2700
	TACGCCATAG	CCGAACGAGC	AGCTCAGAGC	GTTTGTATAT	CATGCTGCTA	2750
5	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	CTTGGCGGCG	TTTATGCTCA	2800
	GAACACAGGT	TGGGACAAGC	ACTTCCAGGC	TAACACAGTC	AGAAATCGAA	2850
	ACGTACTCTC	AACAGTTCCG	TTAGGCATGG	AACTTTTTCG	GCATTCTGCG	2900
	TACACAAATA	CAGGGGAAGA	CTTACTCGTG	GCTGCAACCC	TACTAGCTCA	2950
	AAATTTTATC	ACACATGGTT	AGCGTTTGGG	GAATATTATGA	TAATGATCCA	3000
10	GATCATTCTT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTGTGTTGG	3050
	TTTTTTGTGG	ATCTGCTGTG	CCTTCTAGTT	GCCAGCCATC	TGTTGTTTGC	3100
	CCCTCCCGCG	TGCCTTCTCT	GACCTTGGAA	GGTGCCACTC	CCACTGTCTT	3150
	TTCTCAATAA	AATGAGGAAA	TTGCATCGCA	TTGTCTGAGT	AGGTGTCAAT	3200
	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACA	GCAAGGGGGA	GGATTGGGAA	3250
15	GACAAATGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAGG	TGCTGAAGAA	3400
	TTGACCCGGT	GACCAAGAGT	GCCTTTTATC	ATCACTTTAA	AAATAAAAAA	3450
	CAATTACTCA	GTGCTCTGTA	TAAGCAGCAA	TTAATTATGA	TTGATGCTCA	3500
20	CATCACAACA	AAACTGTGAT	TAACAAATGG	TTGGTCTGCC	TTAGAAAATA	3550
	TATTGAACA	TTATCTTGAT	TATATTATTG	ATAATAATAA	AAACCTTATC	3600
	CCATGCCAAG	AAGTGATGCC	TATCATTTGG	TGGAATGAAC	TTGAAAAAAA	3650
	TAGGCTTGGA	ATACATTACT	GGTAAGGTAA	ACGCCATTGT	CAGCAAAATG	3700
	ATCCAGAGAG	ACCAACTTAA	AGCTTTCTGT	ACGGAAATGT	AATTTCTGTT	3750
25	GACCCTTGAG	ACTGATGAAT	CCCCTAATGA	TTTTTGGTAA	AATCTATTAG	3800
	TTAAGTGGGA	TACACATCTT	GTCAATATAT	CCCGGTAAATG	TGAGTTTCTG	3850
	CATCTATTAG	GCACCCGAGG	CTTTACACTT	TATGCTTCCG	GCTCGTATGT	3900
	TGTTGGGAAT	TGTGGCGGGA	TAACAATTTC	ACACAGGAAA	CAGCTATGAC	3950
30	CATGATTAGC	CCAAGCGCGC	AATTAACCTT	CACTAAAGGG	AACAAGAACT	4000
	GGAGCTCCAC	CGCGGTGGCG	GCCGCTTAGT	AACTAGTGGG	TCCCGCGGGG	4050
	AGGTCAAGAT	GGTTTCTTTA	CTGTTTGTCG	ATTTCTATTAT	TTCAATACAG	4100
	AACAATAGCT	TCTATAACTG	AAATATATTT	GCTATTGTAT	ATTATGATTG	4150
	TCCCTCGAAC	CATGAACACT	CCTCCAGCTG	AATTTACAAA	TTCTCTGTCT	4200
	ATTTGCCGAG	CCATTAAAGT	ATTCAATGAA	GATCTTTTGG	GAACACTGCA	4250
35	AGTTCATATC	ATAAACACAT	TTGAATTGGA	GTATTGTFTT	GCATTGTATG	4300
	GAGCTATGTT	TTGCTGTATC	CTCAGAAAAA	AAGTTTGTGA	TAAAGCATTG	4350
	ACAGCCATAA	AAAGATAGAT	TTAAATATTG	CAGCTATAGG	AAAGAAAGTG	4400
	CGTCTGCTCT	TCACCTTAGT	CTCAGTTGGC	TCCTTCACAT	GCATGCTTCT	4450
40	TATTTCTCCT	TATTTTGTCA	AGAAAAATAT	AGGTCACGTC	TTGTTCTCAC	4500
	TTATGTCTCTG	CCTAGCATGG	CTCAGATGCA	CGTTGTAGAT	ACAAGAAAGG	4550
	TCAAGTAAAA	CAGACTTCTG	GTCTGTTACT	ACAACCATAG	TAATAAGGAC	4600
	ACTAACTAAT	AATTTGCTAAT	TATGTTTTCC	ATCTCTAAGG	TTCCCACTAT	4650
	TTTTCTGTTT	CTTAAAGATG	CCATTATCTG	GTGTAACTG	AACTCTCAATG	4700
	GAACATGAGC	AATATTTCCT	AGTCTTCTCT	CCCATCCAA	AGTCTGTATG	4750
45	GATTAGCAGA	ACAGGCAGAA	AACACATTGT	TACCCAGAA	TAAAACTAA	4800
	TATTTGCTCT	CCATTCAACT	CAAAATGGAC	CTATTGAAAC	TAAAACTTAA	4850
	CCCAATCCCA	TTAAATGATT	TCTATGGCGT	CAAAAGTCAA	ACTTCTGAAG	4900
	GGAACTGTGG	GGTGGGTCAC	AATTCAGGCT	ATATATTCCC	CAGGGCTCAG	4950
50	CGGATCCATG	GGCTCCATCG	GCGCAGCAAG	CATGGAATTT	TGTTTTGATG	5000
	TATTCAAGGA	GCTCAAAGTC	CACCATGGCA	ATGGAACAT	CTTCTACTGC	5050
	CCCATTTGCC	TCAATGTCAG	TCTAGCCATG	GTATACCTGG	GTGCAAAAGA	5100
	CACACCCAGG	ACACGATAAA	ATAAGGTTGT	TGCTTTTATG	AAACTTCCAG	5150
	GATTTCGAGA	CAGTATTGAA	GCTCAGTTGT	GCACATCTGT	AAACGTTTCA	5200
	TCTTCACTTA	GAGACATCCT	CAACCAATTC	ACCAACCCAA	ATGATGTTTA	5250
55	TTCTGTCAGC	TTTGCCAGTA	GACTTTATGC	TGAAGAGAGA	TACCCAATCC	5300
	TGCCAGAATA	CTTGCAAGTG	GTGAAGGAAC	TGTATAGAGG	AGGCTTGGAA	5350
	CCTATCAACT	TTCAACACAG	TGCAGATCAA	GCCACAGAGC	TCATCAATTC	5400
	CTGGGTAGAA	AGTCAGACAA	ATGGAATTTAT	CAGAAATGTC	CTTCAGCCCA	5450
	GCTCCGTGGA	TTCTCAAACT	GCAATGGTTT	TGGTTAATGC	CATTGTCTTC	5500
60	AAAGGACTGT	GGGAGAAAAC	ATTTAAGGAT	GAAGACACAC	AAGCAATGCC	5550
	TTTCAGAGTG	ACTGAGCAAG	AAAGCAAAAC	TGTGCAGATG	ATGTACCAGA	5600

	TTGGTTTATT	TAGAGTGGCA	TCAATGGCCT	CTGAGAAAAA	GAAGATCCTG	5650
	GAGCTTCCAT	TTGCCAGTGG	GACAATGAGC	ATGTTGGTGG	TGTTGCCTGA	5700
	TGAAGTCTCA	GGCCTTGAGC	AGCTTGAGAG	TATAATCAAC	TTTGAAAAAC	5750
	TGACTGAATG	GACCACTTCT	AATGTTATGG	AAGAGAGGAA	GATCAAAGTG	5800
5	TACTTACTCT	GCATGAAGAT	GGAGGAAAAA	TACAACCTCA	CATCTGTCTT	5850
	AATGGCTATG	GGCATTACTG	ACGTGTTTAG	CTCTTCAGCC	AATCTGTCTG	5900
	GCATCTCCTC	AGCAGAGAGC	CTGAAGATAT	CTCAAGCTGT	CCATGCAGCA	5950
	CATGCAGAAA	TCAATGAAGC	AGGCAGAGAG	TGGTAGGGT	CAGCAGAGCC	6000
10	TGGAGTGGAT	GCTGCAGGCG	TCTCTGAAGA	ATTTAGGGCT	GACCATTCAT	6050
	TCCTCTTCTG	TATCAAGCAC	ATCGCAACCA	ACGCCGTTC	CTTCTTTGGC	6100
	AGATGTGTTT	CCCTCCGCG	GCCAGCAGAT	GACGCACCAG	CAGATGACGC	6150
	ACCAGCAGAT	GACGCACCAG	CAGATGACGC	ACCAGCAGAT	GACGCACCAG	6200
	CAGATGACGC	AACAACATGT	ATCTCTGAA	GCTCTTGTGG	CTGGATCGGC	6250
	CTGCTGGATG	ACGATGACAA	ATTTGTGAAC	CAACACCTGT	GCGGCTCACA	6300
15	CCTGTGGGAA	GCTCTCTACC	TAGTGTGCGG	GGAACGAGCC	TCTCTTCACA	6350
	CACCCAGAGC	CCGCGGGGAG	GCAGAGGACC	TGCAGGTGGG	GCAGGTGGAG	6400
	CTGGCGGGGG	GCCCTGTGTC	AGGCAGCCTG	CAGCCCTTGG	CCCTGGAGGG	6450
	GTCCCTGCAG	AAGCGTGGCA	TTGTGGAACT	ATGCTGTACC	AGCATCTGCT	6500
20	CCCTCTACCA	GCTGGAGAAC	TACTGCAACT	AGGCGCCTG	GATCCAGATC	6550
	ACTTCTGGCT	AATAAAGAT	CAGAGCTCTA	GAGATCTGTT	TGTTGGTTTT	6600
	TTGTGGATCT	GCTGTGCCTT	CTAGTTGCCA	GCCATCTGTT	GTTTGGCCCT	6650
	CCCCCGTGCC	TTCTTTGACC	CTGGAGGGTG	CCACTCCAC	TGTCCTTTCC	6700
	TAATAAATAG	AGGAAATTGC	ATCGCATTTG	CTGAGTAGGT	GTCATTTCTAT	6750
	CTGTGGGGGT	GGGTGGGGCG	AGCAACAGCA	GGGGGAGGAT	TGGGAAGACA	6800
25	ATAGCAGGCA	TGCTGGGGAT	GCGGTGGGCT	CTATGGGTAC	CTCTCTCTCT	6850
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACTCT	CTCGAGGGGG	6900
	GGCCCGGTAC	CCAATTGCGC	CTATAGTAGT	TGTTATTACG	CGGCTCACT	6950
	GGCCGCTGTT	TTACAACGTC	GTGACTGGGA	AAACCTTGGC	GTTACCCACA	7000
	TTAATCGCCT	TGCAGCACAT	CCCCCTTTCG	CCAGCTGGCG	TAATAGCGAA	7050
30	GAGGCCCGCA	CCGATCGCCC	TTCCCAACAG	TTGCGCAGCT	TGAATGGCGA	7100
	ATGGAAATG	TAAGCGTTAA	TATTTTGTGA	AAATTCGCGT	TAAATTTTGT	7150
	TTAAATCAGC	TCAATTTTGA	ACCAATAGGC	CGAAATCGGC	AAAATCCCTT	7200
	ATAAATCAAA	AGAATAGACC	GAGATAGGGT	TGAGTGTGTG	TCCAAGTTGG	7250
	AACAAGAGTC	CATATTATAA	GAACGTGGAC	TCCCAACGTC	AAGGGCGAAA	7300
35	AACCGCTCAT	CAGGCGGATG	GCCCACTACT	CCGGGATCAT	ATGACAAGAT	7350
	GTGTATCCAC	CTTAACCTTAA	TGATTTTTAC	CAAAATCAIT	AGGGGATTCA	7400
	TCAGTGCTCA	GGGTCAACGA	GAATTAACAT	TCCGTACAGG	AAGCTTATGA	7450
	TGATGATGTG	CTTAAAAACT	TACTCAATGG	CTGGTTATGC	ATATCGCAAT	7500
	ACATGCGAAA	AACCTAAAG	AGCTTGCCGA	TAAAAAAGGC	CAATTTATTG	7550
40	CTATTTACCG	CGGCTTTTGA	TTGAGCTTGA	AAGATAAATA	AAATAGATAT	7600
	GTTTTTATTG	AAGCTAAATC	TCTTTTATCG	TAAAAAATGC	CCTCTTGGGT	7650
	TATCAAGAGG	GTCATTATAT	TTCCGCGAAT	AACATCATTT	GGTGACGAAA	7700
	TAACTAAGCA	CTTGTCTCCT	GTTTACTCCC	CTGAGCTTGA	GGGGTTAACA	7750
45	TGAAGGTCAT	CGATAGCAGG	ATAATAATAC	ATGAAAAACG	TAAACCAATA	7800
	ATCCAAATCC	AGGCATCCCA	AATTGGTAGT	GAATGATAT	AAATTAACAG	7850
	AAACAGTAAT	GGGCCAATAA	CACCGTTCG	ATTGGTAAGG	CTCAACAAAT	7900
	ATCCCTGTAA	AGCACTTTCG	TGATGACTCT	TGTTTGGAT	AGACATCACT	7950
	CCCTGTAAATG	CAGGTAAGAC	CAGCCCAATA	CAAGCCAAATA	AAATTAACAC	8000
	AGGGAATACT	AACCAACCTT	CAGATATAAA	CGCTAAAAAG	GCAAAATGAC	8050
50	TACTATCTGC	AATAAATCCG	AGCAGTACTG	CGGTTTTTTC	GCCCAATTGA	8100
	GTGGCTATT	TTCTTGCCAC	AAAGCGCTGG	AATACTGAGT	GTAAGAGACC	8150
	AAGACCCGCT	AATGAAAAAG	CAACCATCAT	GCTATTCAT	CCAAAACGAT	8200
	TTTCGGTAAA	TAGCACCCAC	ACCGTTGCGG	GAATTTGGCC	TATCAATTGG	8250
55	GCTGAAAAAT	AAATAATCAA	CAAAATGGCA	TGCTTTTAAA	TAAAGTGATG	8300
	TATACCGAAT	TCAGCTTTTG	TTCCCTTTAG	TGAGGGTTAA	TTGCGGCTCT	8350
	GGCGGTAATCA	TGTCATAGC	TGTTTCTCGT	GTGAAATGT	TATCCGCTCA	8400
	CAATTTCCACA	CAACATACGA	GCCGGAAGCA	TAAAGTGTA	AGCCTGGGGT	8450
	GCCTAATGAG	TGAGCTAACT	CACATTAATT	GCGTTGCGCT	CAGTGGCCGC	8500
	TTTCCAGTGC	GGAACCTGT	CGTCCAGCT	GCATTAATGA	ATCGGCCAAC	8550
60	GCGCGGGGAG	AGGCGGTTTG	GATATTGGGC	GCTCTTCCGC	TTCCTGCTCT	8600
	ACTGACTCGC	TGCGCTCGGT	CGTTCGGCTG	CGCGAGCGG	TATCAGCTCA	8650

CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA 8700
 AGACATGTG AGCAAAAGGC CAGCAAAAGCC CACGGAACCG TAAAAAGGCC 8750
 GCGTGTCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG AGCATCACAA 8800
 AAATCGACGC TCAAGTCAGA GGTGGGGA A CCGCAGGAGA CTATAAGAT 8850
 5 ACCAGCGGTT TCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCGACG 8900
 CTGCGGCTTA CGGGTAACCT GTCCGCCPTT CTCCTTCGG GAAGCGTGGC 8950
 GCTTTCTCAT AGCTCAGGCT GTAGGTATCT CAGTTCGGTG TAGGTCTGTC 9000
 GCTCCAAGCT GGGCTGTGTG CAGCAACCCC CGGTTCAGCC CGACCGCTGC 9050
 10 GCCTTATCCG GTAACATATG TCTTGAGTCC AACCCGGTAA GACGAGACTT 9100
 ATGCCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG 9150
 TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCTTAACTA CGGTACACT 9200
 AGAAGGACAG TATTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG 9250
 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACNAAACCAC GCTGGTAGCG 9300
 GTGGTTTTTT TGTGTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT 9350
 15 CAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAAAGC 9400
 AAATCTCAGT TAAGGGATTT TGGTCTAGAG ATTATCAAAA AGGATCTTCA 9450
 CCTAGATCTT TTTAAATTAA AATGAAGTT TTAATCAAT CTAAAGTATA 9500
 TATGAGTAAA CTTGGTCTGA CAGTTACCAG TGCTTAATCA GTAGGCAACC 9550
 20 TATCTCAGCG ATCTGTCTAT TCTGTTTATC CATAGTTGCC TGACTCCCCG 9600
 TCGTGTAGAT AACTACGATA CGGAGGGCTT TACCATCTGG CCCAGTCTCT 9650
 CGAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT 9700
 AAACCGAGCA CGCGAAGCG CCGAGGCGAG AAGTGGTCTT GCACTTTAT 9750
 CGCGCTCCAT CAGTCTATT AATTGTGTCC GGAAGCTAG AGTAAGTAGT 9800
 25 TCGCCAGTTA ATAGTTTGGC CACCGTTGTT GCCATTGTCT CAGGCATCGT 9850
 GGTGTACGCG TGCTGTGTTG GTATGGCTTC ATTACGCTCC GGTTCCTAAC 9900
 GATCAGGCGG AGTACATGA TCCCCCATGT TGTGCAAAAA AGCGGTTAGC 9950
 TCCCTTGGCT CTCGATCGT TGTCAGAGT AAGTTGGCG CAGTGTATC 10000
 ACTCATGTT ATGGCAGCAC TGCTAATATC TCTTACGTG ATGCCATCCG 10050
 TAAGATGCTT TCTGTGACT GGTGAGTACT CAAACCAAGT ATTCTGAGAA 10100
 30 TATGTATGCG GCGACCGAG TGTCTCTTCG CCGCGCTCAA TAGCGGATAA 10150
 TACCGCGCCA CATAGCAGAA CTTTAAAGT GCTCATCAT GGAACAGCTT 10200
 CTTGGGGGCG AAATCTCTCA AGGATCTTAC CGCTGTGTAG ATCCAGTTGG 10250
 ATGTAACCCA CTCGTGCAAC CAACGTATCT TCAGCATCTT TTACTTTTAC 10300
 CAGCGTTTCT GGTGAGCAA AAACAGGAAG GCAGAAATGCC GCAGAAAGG 10350
 35 GAATAAGGGC GACAACGAAA TGTGTAATAC TCATACTCTT CCTTTTTCAG 10400
 TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGACG GATACATATT 10450
 TGAATGTATT TAGAAAAATA AACAAATAGG GGTCCCGCG ACATTTCGCC 10500
 GAAAGTGCC AC 10512

40
 SEQ ID NO:43 (pTnMOD (CMV-CHOVg-ent-ProInsulin-synPA))

1 ctgacgcgcc ctgtagcgcc gcattaaagc cggcgggtgt ggtgggttac cgcagcggtga
 45 61 cggctaacat tgcacagcgcc ctagcgccgc ctcctttgcg tctcttccct tccctttctg
 121 caagcttgcg cgcacataga ctggctattg gccattgcgt acgtttgtac catatacata
 181 tatgtacatt tatattggct cabgctcaac attacgcgca tgttgacatt gattattgac
 241 tatgtataa tegttaataa ctacgggtgt attagtcctat agcccatata tggagttccg
 301 cgttaacata cttaacgttaa atggcgccgc tggctgacgc cccaaagccc ccgcgccatt
 361 gaagtaata atgcgtatg tcccatagc aacgcataa gggactttcc atgaogtca
 421 atgggtgagc tatttcgggt aaactgccc ctgggaagta catcaaggtt atcatatgcc
 481 aagtagccc cctattggcg tcaatgaagg taaatggccc gectggcatt attgccagta
 541 catgaactta tgggactttc ctacttggca gtacatcac gattatgata tgcattatgc
 601 catggtgatg cgtttttggc agtacatcaa tggcggtgga tagcggtttg actcaagggg
 661 atttccnagt ctccaccgca ttgaogtcaa tgggtttgtt ttctggcacc aaatcaagc
 721 gaaacttcca aaatgtcgta acaactccgc cccatagaag acacgggagc gcatacagc tccggcgccg
 781 acggtggagc gtcatataa cgaagagctg tttagtgaac cgtacagatt cctggagagc
 841 ccatccagcg ttttttgacc tccatagaag acacgggagc gcatacagc tccggcgccg
 901 gaaacggtgc attggaagc ggaattccgc tgcacaggtg gacgtatga cgcctatagc
 961 actcatagc caacccctt tggctcttat gcatgtcata ctgtttttgg ctgtggggct
 1021 atacacccc gcttcttat gctataggtg atggtatagc ttacgctata ggtgtgggtt
 1081 attgacatt attgaacact cccattagg tgaogatac tctatatt aatccataac
 1141 atggctcttt gccacacata tctctattg ctatagcca atactctgtc cttcagagac
 1201 tgcacggagc ctgtatttt tacagatgg ggtccattt attatttaca aattcacata

	1261	tacaacaacg	cogtccoccg	tgcocgcagt	ttttattaaa	catagcgtgg	gatctccacg
	1321	cgaaatctcgg	gtacgtgttc	cggaacatggg	ctctctcgcg	gtagcggcgg	agcttcacaa
	1381	tcogagccct	ggtoaccatgc	ctccagcgggc	toatggctgc	tcggcagctc	cttgctccaa
5	1441	aagcgtggagg	ccagacttag	gcacagcaca	atgccaccca	ccaccagctg	gcgcgacaa
	1501	gcgcgtggcg	tagggatatgt	gtctgaaaaat	gagcgtggag	atggggctcg	caocgctgac
	1561	gcagatgggaa	gacttaaggc	agcggcagaa	gaagatgcag	gcagctgagt	tgtgttatcc
	1621	tgataagagat	caagaggtaac	tcocgcttgcg	gtcgtgttaa	cggctggagg	cagtgtagtc
	1681	tgacagctac	tcgttgtctgc	cgocgocgoc	acacagacata	atagctgcaca	gactaacagaa
10	1741	ctgttcccttt	ccatggggtct	ttctctcagt	caocgctoga	ccatgtgtga	acttgabatt
	1801	ttacatgatt	ctctttaacca	attctgcccc	gaattacaact	taaaacagt	caacagctta
	1861	acgttggcttt	gcacgcattt	acttgactgt	aaaaactotca	ctcttaacga	acttgccgct
	1921	aacctgcocaa	ccaaagcgag	acaaaacat	aacatacaaac	gaatgcacg	atttgttagt
	1981	aacgtgtcaac	tcacaaaaga	ggacttcgct	gtatacogtt	ggacgtgctg	cttgttatgt
15	2041	tcggcgaaata	cgatgcccat	tgtacttgttt	gactggctgtg	ataticgtga	gcacaaacga
	2101	ctattgttat	tcgagacttc	agtcgcacta	caecggttgt	ctgtcaact	tatatgagaa
	2161	gogtccccgc	tttcagagca	agtttcaaa	aaagctcact	accaatttct	agccgaactt
	2221	cgagagcattc	taocgagtaa	caccaacgc	ctcaattgca	gtgatgtcgt	ctttaaattg
	2281	tcattgttata	aatccgttga	gaagctgggt	tggtaactgt	taagtccagt	aagaggaaaa
20	2341	gtacaaatgt	cagacatagg	agcggaaaa	tggaaaacct	tcagcaact	actgatattg
	2401	ctcatcagtc	actcaaacat	tttagctat	aaagggctga	ctaaagcga	tcocactcca
	2461	tgccaaatttc	tattgtataa	atctcgtctc	aaagggcgaa	aaaatcagcg	ctgcacagcc
	2521	actcaatgtc	accacccgtc	actcaaaatc	tactcagctg	cgccgaagga	gcacatgggtt
	2581	ctagcaacta	acttaoctgt	tgaaatttga	acacccaaac	acttgtttaa	cttattttcg
	2641	agcattgaga	agattgaaga	acacttccga	gaacttgaaa	gtcctgccta	cggaatagcg
25	2701	ctacgcaata	gcgcgaacag	cagctcagag	cggtttgata	tactgctgc	atcgcocctc
	2761	atgcttcaac	taacatgttg	gcttgccggc	gttctatgct	agaaacaagg	ttgggacaa
	2821	cacttccagg	ctaacacagt	cagaaatcga	aacgtactct	caacagttcg	cttgatattg
	2881	gaagttttgc	ggcattctcg	ctacacaaata	acaaaggga	acttaactcg	ggctgcacac
	2941	ctactagctc	aaaattttat	cacacatgg	tacgcttttg	ggaaattat	ataatgatcc
30	3001	agatcaactc	tggttaataa	aagatcagag	ctctagagat	ctgtgtgttg	gttttttgtg
	3061	gatctgtgtg	tccttctagt	tgccagccat	ctgtgtgttg	gcctccccc	gtgccttctc
	3121	tgacccctga	aggtgcccct	cccactgtcc	tttctcaata	aaatgaggaa	attgatcgc
	3181	attgtcttga	taggtgtcat	ttctattctg	gggtgggggt	ggggcagcac	agcgaagggg
35	3241	aggatctggg	agcaaatagc	aggcatgctg	gggtccgggt	ggcctctatg	ggctactctc
	3301	ttctctctct	ttctctctct	ttctctctct	ttctctctct	ttctctctct	ttctctctct
	3361	ttctctctct	ttctctctct	tcggtaccag	gtcgtgaaga	attgaccccg	tgacaaaagg
	3421	tgcccttttat	catcaacttta	aaaataaaaa	acaattactc	agtgccctgt	ataagcaga
	3481	attaattatg	attgatgctc	acatcaaac	aaaaactgat	ttacaacatg	gttggtctgc
40	3541	atttgaaagt	atatttgaac	attatcttga	ttatatatta	gtataataa	aaaacactta
	3601	cccatatcca	gaagtgatgc	ctatcattgg	ttggaaatga	cttgaaaaaa	attagccctg
	3661	aatacaattac	tggtgaaggt	aacgcacattg	tcagcaaatg	gatccaagag	acccaactta
	3721	aagctttcct	gaaggaatgt	taattctcgt	tgacccctgag	caactgatga	tcocctaatg
	3781	attttgtgaa	aaatcattaa	gttaaggctg	atacacatct	tgctcatatg	tcocggtaat
45	3841	gtgagttagc	tcactcatta	ggcacccacg	gctttacact	ttatgcttcc	ggctcgtatg
	3901	ttgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acagctatga	ccatgatatac
	3961	gccaaagcgc	caattaaccc	tcactaaagg	gaacaaaagg	tgagctcca	cgcgggtggc
	4021	ggccgcctcta	gaactagtgc	atcccccggg	catcagatgt	gctattggcc	attgatcaag
	4081	ttgtatccat	atcataatat	gtacacttat	atttgctcat	gtccaacatt	accgcacatg
	4141	tgacactgat	tattgactag	ttattaatag	taatcaattt	cgggctcatt	agttctatgc
50	4201	ccatatatgg	agttcccggt	taacataact	acggtaaatg	gcocgcctgg	ctgacgcgcc
	4261	aagacccccc	gcccattgac	gtcaataatg	agctatgttc	catagtaac	gcccaatagg
	4321	actttccatt	gacgtcaatg	gggtggagta	ttacgggtaaa	ctgcccaact	ggcgactcat
	4381	caagtgatc	atatgccaa	tacgccccct	atgacgctga	atgacggtaa	atggcccgcc
55	4441	tgccgatgat	ccagtgatct	gaccttatgg	gactttcccta	cttgccagta	catctaatga
	4501	ttagtcatcg	ctattacocat	ggatgatcgg	ttttggcagt	acatcaatgg	ggctggatag
	4561	cggtttgact	caoggggatt	tcocaaagtct	caccocattg	agctcaatgg	gagtttgttt
	4621	tggtcccaaa	atcaacggga	ctttccaaaa	ttgtcgttaaca	actccgcgcc	ctgacgcgaa
	4681	atggggcggt	ggcgtgtgac	gtggggaggtc	tatatagaoca	gagctcgttt	agtgaaacct
60	4741	cagatcgctc	ggagacgcca	tcacgcctgt	tttgacctcc	atagaaagca	cggggaacga
	4801	tcagacccct	ggggccggga	acgtgtcatt	ggaaocggga	ttccocctgc	caagatgacg
	4861	gtaagtaacg	ccctatagact	ctataggcac	acocctttgg	ctcttatgca	gtctatactg
	4921	tttttggctt	ggggccctata	cacccccgct	tcctttatgt	atagggtgat	gtatagctta
	4981	gcctatagg	gtgggttatt	gaaccaattt	gaaccaactcc	ctattgggtga	gaactacctt
65	5041	ctactcaaat	ccataacatg	gctctttgct	acaactatct	ctattgggtga	tcgtccaata
	5101	ctctgtccct	ccagagactga	cacggagctcc	gtattttttac	aggaatgggt	ccattttatt
	5161	atttacaatt	tcacataatc	acacaaagcgc	tcocccgtgc	cgcagttttt	ttttaaacat
	5221	agcgtgggat	ctccaaogga	actcggggta	cgtgttcggg	acaatgggctc	ctctccgcta
	5281	ggcggggagg	ttccacatcc	gagccctgggt	cccatgcctc	cagcggctca	tggtgcctcg

	5341	gcagctcctt	gctcctaaca	gtggaggcca	gacttaggca	cagcacaaatg	cccacaccca
	5401	ccagtggtgc	gcacaaggcc	gtggcggtag	gggtatgtgtc	tgaaaatbag	cggtggagatt
	5461	gggctgcgac	ggctgcagcca	gatggaagac	ttaaaggcagc	ggcagaagaag	gatgcaggca
	5521	gctgagttgt	tgtattctga	taagagtcag	aggtaactcc	cgttgcggtg	ctgttaaagg
5	5581	tgaggggcag	tgtagtctga	gcagtaactcg	ttctgcgcgc	ggcgcgcaacc	agacataata
	5641	gctgcagac	taacgacagt	ttcctttcca	ttgggtctttt	ctcgagtcac	cgtcggaaca
	5701	atggggctcca	tcgggtgcagc	aagcatbgaa	ttttgttttg	atgtattcaa	ggagctcaca
	5761	gtccacacatg	ccaatgagaa	catctttcac	tgccccatgt	ccatcatgtc	cggtgataca
	5821	atggatatacc	tggttgcaaaa	agacagcaccc	aggacacaaa	taataaaggt	tgctcgcttt
10	5881	gatacaacttc	caggattccg	agacagattc	gatgcactag	gtggcacaalc	tgtaaacgtt
	5941	caactcttcac	ttagagacat	ctccaaccaa	atccacaacac	caaaatgatgt	ttatcgcttc
	6001	agcttgtcca	gtagacttta	tgcgtgaagag	agatacccaa	tcctgcagaa	ataacttgac
	6061	tggtgtgaagg	aactgttatag	aggaaggtctg	gaacctatca	actttcaaac	agctgcagat
	6121	cacgcacagag	agctcatcaaa	ttcctgggta	gaagctcaga	caaatggaaat	tatcagaat
15	6181	gtcctctcagc	caagctccgt	ggattctccaa	actgcaatgg	ttctgtgttaa	tgcoattgtc
	6241	ttcaaaaggac	tgtgggagaa	agcattttaag	gatgaagacc	cacaagcaat	goccttcaga
	6301	gtgactgagc	aagaaaagcaa	acctgtgcag	atgatgtacc	cattttgcag	tgggacaaatg
	6361	gcataaatgg	ctctctgagaa	aatgaagatc	ctagagcttc	agcagcttga	tgataatac
	6421	agcattgttg	tgtgtgtgoc	tgaagaagtc	tcagcgcttc	cattttgcag	tggaacaaatg
20	6481	taacttgtaaa	aactgactga	atggagccagt	taatacaact	caactctgtc	taaatggcta
	6541	tgtactatacc	tgcaatgaag	tgacgtgttt	tctaagtta	tggaagagag	aagatacaag
	6601	tggggaattac	tgacgtgttt	agctcttcag	caaatgtctc	tggaactctcc	tcacgagaga
	6661	gcttgaagat	atcttaagct	gtccatggaag	caactgcaga	aatcaatgca	gacggagag
	6721	agggtggtagg	gtccagcagag	gctggaggtg	atgcctgaag	cgctctgaa	gaattatagg
25	6781	ctgacacattg	attctctctc	tgtatcaagc	acatgcgaac	caacgcgctt	ctctctcttt
	6841	ggcagatgtg	tttcccgagg	ccagcagatg	acgcacacagc	agatgacgca	ccagcagatg
	6901	agccacacag	agatgacgca	ccagcagatg	acgcacacac	atgtatcctg	aaagctcttt
	6961	gtggctggat	cggtctgctg	gatgacgatg	acaaatttgt	gaaccaaacac	ctgtcgggct
	7021	cacacactgt	ggagactctc	taactatgtg	cgcggggaagc	aggtctcttc	tacacaacca
30	7081	agaccccgcg	ggagggcagag	gaactgcagc	tgggggcaggt	ggagctgggc	ggggggccctg
	7141	gtgcagcgag	ctgcgcagcc	ttggccctgg	aggggtccct	cgagaagagc	ggcatctgtg
	7201	aaacatgctg	taccagcatc	tgtctctctc	accagctgga	gaactactgc	aactagggct
	7261	ctctaagggc	gaattatcgc	ggccgctcta	gaccaggcg	ctggatccag	ctactctctg
	7321	gctaataaaa	gatcagagct	ctagagatct	gtgtgttggt	tttttgtgga	tcgtgtgtgc
35	7381	ctctctagtgt	ccagccatct	gttctttgcc	ctcccccgtg	ctctctctg	acccttgagc
	7441	gtgcacatcc	cactgtctct	tctataaaaa	atggaggaaat	tgaatcgcat	tgtctgata
	7501	ggtgtcattc	tattctgggg	ggtgggggtg	ggcagcacag	caagggggag	gattgggaag
	7561	acaatagcag	gcattgtggg	gatgcgggtg	gtctctatgg	taactctctc	tctctctctc
	7621	ttctctctac	tctctctctc	tctctgtacc	ttctctcgag	ggggggcccg	gtacccaatt
40	7681	cgccctatag	tgagtgtgat	tacgcgcgct	caactgtccg	cggtttacaa	cgctgtgact
	7741	tggaataaacc	tgccgttacc	caacttaact	gccttgcaag	acactcccc	ttgcgcagct
	7801	ggcgtaatatg	cgaagagggc	cgcacgcgct	gccttcccca	acagttggcg	agcctgaatg
	7861	gcgaatggaa	attgtaagcg	ttaattatttt	gttaaaattc	gcgttaaat	ttgttgtaat
	7921	cagctcattt	tttaaccaat	aggcogaagt	cggaacaaatc	ctctataaat	caaaagaata
45	7981	gaaccagata	gggttgagtg	ttgttccagt	ttggaacag	agtcacact	taactccaag
	8041	ggactcccaac	gtcaaaaggcg	gaaaaacccg	ctatcagggc	gatggccac	tactccggga
	8101	tcatatgcaca	agatgtgtat	ccaccttaac	ttaatgattt	ttacaaaaat	ctatagggaa
	8161	ttcatcagtg	ctcaggggtca	acgagaattat	acattccctg	aggaagagct	atgatgata
	8221	tgtgtctaaa	aacttaactca	atggctgggt	atgcatatcg	caatacatgc	gaaaaaacct
50	8281	aaagagcctg	cogataaaaa	agggccaatt	atgtctattt	accccggtt	tttatgagc
	8341	ttgaaagata	aataaaaatg	atagggttta	tttgaagcta	aatctctctt	atcgtaaaaa
	8401	atgccctctt	gggttatcaa	gagggctcatt	atatttcgog	gaataacatc	atttggtgac
	8461	gaataaacta	agcaactgtc	ttctgtttac	ttccctgagc	ttgaggggtg	aacatgaag
55	8521	ctatcgatag	caggataata	atacagttaa	acgctaaacc	ataataccaa	atcccgacat
	8581	ccocaaatgg	tagtgaatga	ttataaataa	cagcaaaacg	taatggggca	ataaacccg
	8641	tgtcatttgt	aaggctcacc	aataatccct	gtaaagacac	ttgtgatga	ctcttggtt
	8701	ggatagacat	cactccctgt	aatgcaggta	aagcagatcc	accaccagcc	ctatgaatga
	8761	aaacagggaa	aactaaccaa	cttcagata	taaaogctta	aaaggcaaat	gcactcctat
	8821	ctgcataaaa	tcagagcagt	actgcgcttt	tttgcgcat	ttagtggata	ttctctcgc
60	8881	cacaaaggct	tggaataactg	agtgtaaaag	accgaacccc	gtaatgaaaa	ggccacatc
	8941	atgctatcca	tcactcagat	ctctgttaata	gcacacaccc	gtgtggatc	cgatcgaatt
	9001	gogctgaaat	aataatcaac	aaatggcact	gttaataaag	tgatgtatac	agctgaagtt
	9061	tgttctccctt	tagtggaggt	taattgcgog	ctggcgctaa	taactggcat	agctgtttcc
	9121	tgtgtgaaat	tgttatccgc	tcacaattcc	acacaaacata	cgagccgaaa	gcataaagtt
65	9181	taaagccctg	gggtgcataat	gagtgaagta	actcacaatta	attgcgttgc	ctactcgcg
	9241	cgcttctccag	tcgggaiaacc	tgtcgtgcac	gtgcgataaa	tgaaatcgcc	aaagccggg
	9301	gagagggggt	tgcgttatgt	gggcctctc	cgctctctcg	ctcaactgac	cgctgcgctc
	9361	ggtcgttcgg	ctcggggag	cggatcaagc	tcactcaaaag	gggttaaac	ggttatccac

	9421	agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa
	9481	cogtaaaaag	gcgcggttgc	tgccgttttt	ccaataggctc	cgccccctgc	acagacatca
	9541	caaaaaatcga	cgctcaagtc	agaggtggcg	aaaccgcgaca	ggactataaa	gataccaggc
5	9601	gtttccccc	ggaagctccc	tcgtgtccgc	tcctgttccg	acctgcgcgc	ttaccggata
	9661	cctgtccgcg	ttttccccc	cggaagcgt	ggcgctttct	catagctcac	gctgtaggta
	9721	tctcagttcg	gtgtaggtcg	ttcgtcccaa	gctgggctgt	gtcgacgcac	cccccgctca
	9781	gccgcacgcg	tgccgcttat	ccggttaacta	tcgtcttgag	tcacaaccgg	taagacacga
	9841	cttatcgcca	ctggcagcag	ccactggtaa	caggatagc	agagcgaggt	atgtagccgg
10	9901	tgctcacagag	ttcttgaagt	gggtggcctaa	ccatggctac	actagaagga	cagttattgg
	9961	tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaaga	gttggtagct	ctcagtcggg
	10021	caaaacaacc	accgctggta	gcggctggttt	ttttgtttgc	aagcagcaga	ttacgcgcag
	10081	aaaaaaaagg	tctcaagaag	atcctttgat	cttttctacg	gggtctgaag	ctcagtcggg
	10141	cgaaaaatcca	cgtaaaagga	ttttgtgcat	gagattatca	aaaaaggtat	taacttagat
	10201	ctttttaaact	taaaaatgaa	gttttaactc	aatctaaagt	atactagagt	aaacttggct
15	10261	tgcagcttgc	caatgcttaa	tcagttaggc	acctatctca	gcgatctgtc	tatttcgttc
	10321	atccatagtt	gctgactccc	cgcgtgactcc	gactogtgta	atacggggagg	gcttaccat
	10381	tggccacagc	gctgaagtaa	tacccgcgaga	ccacagctca	ccggctccag	atttatcagc
	10441	aaaaaacagc	ccgcgcggaa	gggcgcggagc	gagaagtggg	cctgcacact	tatccgctcc
	10501	catccagctc	attaagtgtt	gcgcggagct	tcagagtaag	agttccgcgt	ttaatagttt
20	10561	gcgcgaagct	gttgccattg	ctacaggcat	cgttggttca	cgctcgtcgt	ttgtagtgag
	10621	ttcatccagc	ccggttccc	aacgataaag	ggaggttaac	tgatcccoca	tggttgcaa
	10681	aaaaagcgtt	agctcctcgc	gtcctccgat	cgttgtoaga	agtaagttgg	ccgcagtggt
	10741	atactcatg	gttatggcag	cactgcataa	ttctcttact	gctcatggat	cgcctaaggt
25	10801	ctttttctgtg	actgggtgag	actcaaccaa	gtcattctga	gaatagtcta	tgccggcaga
	10861	gagttgctct	tgcccgccgt	caatacggga	taataccgcg	ccacatagca	gtaacttaaa
	10921	agtgctcatc	attggaanaa	gttcttcggg	ggaaaaaactc	ccaaggtatc	taccgcttgt
	10981	gagatccagt	tcgatgtaac	ccactcgtgc	accocaactga	tcttcagcat	ctttactctt
	11041	ccacagcgtt	tctgggtgag	caaaaaacagg	aaggcaaaat	ggcccaaaaa	aggggaatga
	11101	ggcgacaacg	aaatgttgaa	tactcatact	ctctcttttt	caattattat	gaagcatagg
30	11161	tcagggttat	tgtctcatga	gcggatacat	atttgaaatg	atttagaaaa	ataaacaaat
	11221	aggggttcgc	cgcacatttc	ccgaaaagt	gcccac		

	SEQ ID NO:44 (pTnMod(Oval/ENT tag/Proins/PA) - QUAIL)	
35	CTGACGCGGTG CTTGACCGGC GCATTAAAGC CGCGCGCGGT GGTGGTTACG	50
	CGCAGCGTGA CCGGTACACT TGCCAGCGCC CTAGCGCCCC CTCCTTTCGC	100
	TTTCTTCCTT TCCTTTTCCT CCACGTTTCG CGGCATCAGA TTGGCTATTG	150
	GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT	200
	CATGTCCAAC ATTACGCCCA TGTGTGACAT GATTATTGAC TAGTTATTAA	250
40	TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCGG	300
	CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGGTGACCG CCCAACGACC	350
	CCCGCCCACT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA	400
	GGGACCTTTC ATTGACGTCA ATGGGTGGAG TATTACCGGT AACCTGCCAA	450
	CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG	500
45	TCAATGACGG TAAATGGCCC GCCTGGCAAT ATGCCCACTA CATGACCTTA	550
	TGGGACTTTC CTACTTGCCA GTACATCTAC GTATTAGTCA TCGCTATTAC	600
	CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGGCTGGA TAGCGGTTTG	650
	ATCTACCGGG ATTTCCAAGT TCCACCCCA TTAGCGTCAA TGGGAGTTTG	700
50	TTTTGGGACC AAAATCAACG GGACTTTCCA AATAGTCGTA ACAACTCCGC	750
	CCCATTTGAG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA	800
	GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCATCCACGC CCATCCACGC	850
	TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCCGGGCCG	900
	GGAGACGGTC ATTGGAAACG GGATTCCCGG TGCCAAGAGT GACGTAACTA	950
	CGCGCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA	1000
55	CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTTCTTTAT GCTATTAGGT	1050
	ATGGTATTAG TTAGCCCTATA GGTGTGGGTT ATTGACCATT ATTGACCATT	1100
	CCCCATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT	1150
	GCCACACTA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTCAGAGAC	1200
	TGACACGGAC TCTGTATTTT TACAGGATGG GTTCCCAATT ATTATTATGA	1250
60	AATTCCACATA TACAACAACG CGCTCCCCCG TGCCCGCAGT TTTTATTAAA	1300
	CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTC CGGACATGGG	1350
	CTCTTCTCCG GTAGCGGGCG AGCTTCCACA TCCGAGCCCT GGTCCTCATC	1400
	CTCCAGCGGC TCATGTGTCG TCGGCAGCTC CTGTCTCTTA ACAGTGGGAG	1450
	CCAGACTTAG GCACAGCACA ATGCCACCA CCACAGTGT GCCGCACAG	1500

	GCCGTGGCGG	TAGGGTATGT	GTCGTAAAT	GAGCGTGGAG	ATTGGGCTCG	1550
	CAOGGCTGAC	GCAGATGGAA	GACTTAAGGC	AGCGGCGAGA	GAAGATGCAG	1600
	GCAGCTGAGT	TGTTGTATTC	TGATAAGAGT	CAGAGGTAAAC	TCCCGTTGGC	1650
	GTGCTGTAA	CGGTGGAGGG	CAGTGTAGTC	TGAGCAGTAC	TGTTGTCTGC	1700
5	CGCGCGCGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCCCTTT	1750
	CCATGGGTCT	TTTCTGCAGT	CACCGTCGGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGATT	CTCTTTACCA	ATTCTGCCCC	GAATTACACT	TAAAACGACT	1850
	CAACACGCTTA	ACGTTGGCTT	GCCACGCAAT	ACTTGACTGT	AAAACCTCTCA	1900
10	CTCTTACCGA	ACTTGGCCGT	AACTTGCCAA	CCAAAGCGAG	AACAAAACAT	1950
	AACATCAAAAC	GAATCGACCG	ATTGTATAGT	AATCGTCACC	TCCACAAGAA	2000
	GGCACTGCCT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TGCGGAATAC	2050
	GATGCCCATT	GTACTTGTGT	ACTGGCTTGA	TATTCGTGAG	CAAAAACGAC	2100
	TTATGTGATT	GGGAGCTTCA	GTGCACTACT	ACGGTGTCTC	TGTTACTCTT	2150
	TATGAGAAAG	CGTTCCCGCT	TTGAGAGCAA	TGTTCAAAGA	AAGCTCATGA	2200
15	CCAATTTCTA	GCCGACCTTG	CGAGCACTCT	ACCGAGTAAC	ACCACACGCA	2250
	TCAATTGTCAG	TGATGCTGGC	TTTAAAGTGC	CATGGTATAA	ATCCGTTGAG	2300
	AAGCTGGGTT	GGTACTGGTT	AAGTCGAGTA	AGAGGAAAAG	TACAATATGC	2350
	AGACCTTAGGA	GCGGAAAACCT	GGAAACCTPAT	CAGCAACTTA	CATGATATGT	2400
	CATCTAGTCA	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAAGCAAT	2450
20	CCATCTCAT	GCCAAATTTCT	ATTGTATATA	TCTCGCTCTA	AAGGCCCGAAA	2500
	AAATCAGCGC	TGACACGGGA	CTCATTTGTC	CCACCCGTCA	CCTAAAATCT	2550
	ACTCAGCGTC	GGCAAGAGGAG	CCATGGGTTG	TAGCAACTAA	CTTACTCTTT	2600
	GAAATTCGAA	CACCCAAACA	ACTTGTATAT	ATCTATTCTGA	AGCGAATGCA	2650
	TATGAAGAA	ACCTTCCGAG	ACTTGAAGAA	TCTGCTCTAC	GGACTAGGCC	2700
25	TACGCACTAG	CGGAACGAGC	AGCTCAGAGC	GTTTGTATAT	CATGCTGTCTA	2750
	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	TCTGCGGCGC	TTCACTGTCA	2800
	GAACAAGGT	TGGGACAAGC	ACTTCCAGGC	TAAACACAGTC	AGAAATTCGA	2850
	ACGACTCTTC	AACAGTTCGC	TTAGGCATGG	AAGTTTTCGC	GCATTCTGGC	2900
	TACACAATAA	CAAGGGGAAGA	CTTACTCTGT	GCTGCAACCC	TACTAGCTCA	2950
30	AAATTTATTC	ACACATGGTT	ACGCTTTGGG	GAATATAGA	TATGATCCCA	3000
	GATCACTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTTGTGTTG	3050
	TTTTTTTGGG	ATCTGCTGTG	CCTTCTAGTT	GCCAGCCATC	TGTTGTGTTG	3100
	CCCTCCCGCG	TGCTTTCTCT	GACCCCTGAA	GGTGCCACTC	CCACTGTCTC	3150
	TTCTTAATAA	AATGAGGAAA	TTGCATCGCA	TTGTCTGAGT	AGGTGTCTAT	3200
35	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACA	GCAAGGGGGA	GGATTGGGAA	3250
	GACAATAGCA	GGCATGCTGG	GGATGCGGGT	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAGG	TGCTGAAGAA	3400
	TTGACCCGGT	GACCAAAAGT	GCCTTTTATC	ATCACTTTAA	AAATAAAAAA	3450
40	CAATTACTCA	GTGCCCTGTA	TAAGCAGCAA	TTAATTATGA	TTGATGCTCA	3500
	CATCACACAA	AAAACCTGAT	TAACAAATGG	TTGGTCTGCG	TTAGAAAGTA	3550
	TATTTGAACA	TTATCTTGAT	TATATTATTT	ATAATAATAA	AAACCTTATC	3600
	CCATATCAAG	AAGTGATGCC	TATCATTTGT	TGGAATGAAC	TTGAAAAAAA	3650
	TTAGCCCTGA	ATACATTACT	GGTAAGGTAA	AGCCCATTTG	CAGCAAAATTG	3700
45	ATCCAGAGGA	ACCAACTTAA	AGCTTTCCCTG	ACGGAAATGTT	AATTTCTGTT	3750
	GACCCTGAGC	ACTGATGAAT	CCCCTAATGA	TTTTGGTAAA	AATCATTAAG	3800
	TTAAGTGGA	TACACATCTT	GTCAATATGAT	CCCCGTAATG	TGAGTTAGCT	3850
	CACTCATTAG	GCACCCGAGG	CTTTACACTT	TATGCTTCGG	GCTCGTATGT	3900
	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC	3950
50	CATGATTACG	CCAAGCGCGC	AATTAACCTT	CACATAAAGG	AACAAAAGCT	4000
	GGAGCTCCAC	CGCGGTGGCG	GCGGCTCTAG	AACATAGTGA	TCCCCCGGGG	4050
	AGGTCAGAAAT	GGTTTCTTTA	CTGTTTGTC	ATTTCTATTAT	TTCAATACAG	4100
	AACAAAAGCT	TCTATAACTG	AAATATATTT	GCTATTGTAT	ATTATGATTG	4150
	TCCCTCGAAC	CATGAACACT	CCTCCAGCTG	AATTTACAAA	TTCCCTCTGC	4200
55	ATTGTCGAGG	CTGGAAGATC	ATGGAAGATC	TCTGAGGAAC	ATTGCAAGTT	4250
	CATACCATAA	ACTCAATTGG	AATTGAGTAT	TATTTTGCTT	TGAATTGAGC	4300
	TATGTTTTCG	AGTTCCCTCA	GAAGAAAAGC	TGTGTTATAA	GCGCTCTACAC	4350
	CCATCAAAAG	ATATATTTAA	ATATTTCCAAC	TACGAAAAGA	TTTTGTCTGC	4400
	CTTCACTCT	GATCTCAGTT	GGTTTCTTCA	GCTACATGCT	TCITTATTGTT	4450
60	CCTATTTTGT	CAAGAAAATA	ATAGGTCAAG	TCCTGTTCTC	ACTTATCTCC	4500
	TGCCTAGCAT	GGCTTAGATG	CACGTTGTAC	ATTCAAGAG	GATCAAAATGA	4550

	AACAGACTTC	TGGTCTGTTA	CAACAAACCAT	AGTAATAAACC	AGACTAACTA	4600
	ATAATGCTTA	ATTATGTTTT	CCATCTCTAA	GGTTCACCACA	TTTTCTGTGT	4650
	TTAAGATCCC	ATTATCTGGT	TGTAACCTGAA	GCTCAATGGA	ACATGAACAG	4700
	TATTTCTCAG	TCTTTCTCC	AGCAATCCCTG	ACGGATTAGA	AGAAGCTGGCA	4750
5	GAAGAACACTT	TGTTACCCAG	AATTAATAAAC	TAATATTGTG	TCTCCCTTCA	4800
	ATCCAAATG	GACCTATTGA	AACATAAATC	TGACCCAAATC	CCATTAAATT	4850
	ATTTCTATGG	CGTCAAAGGT	CAAACTTTTG	AAGGGAACCT	GTGGGTGGGT	4900
	CCCAATTCAG	GCTATATATT	COCCAGGGCT	CAGCCAGTGG	ATCCATGGCT	4950
	TCCATCGGTG	CAGCAAGCAT	GGAAITTTGT	TTTGATGTAT	TCAAGGAGCT	5000
10	TAAGATCCAC	CATGCCAATG	ACACATGCT	CTACTCCCC	TTTGCCATCT	5050
	CGTCAACTCT	GGCCATGGTC	TTCTAGGTG	CAAAAGACAG	CACCAAGACC	5100
	CAGATAAATA	AGGTGTGTCA	CTTTGATAAA	CTCCAGGAT	TCGGAGACAG	5150
	TATTGAAGCT	CAGTGTGGCA	CATCTGTAAA	TGTTCACTCT	TCACTTAGAT	5200
	ACATATCTCA	CCAAATCACC	AAACAAAATG	ATGCTTATTC	GTTCAAGCTT	5250
	CGCAGTAGAC	TTTATGCTCA	AGAGACATAC	ACAGTCGTGC	CGGAATACCT	5300
15	GCAATGTGTG	AGGGAATGCT	ATAGAGGAGG	CTTAGAATCC	GTCACCTTTC	5350
	AAACAGCTGC	AGATCAAGCC	AGAGGCCTCA	TCAATGCCTG	GGTAGAAGCT	5400
	CAGACAAACG	GAATATATCAG	AAACATCCTT	CAGCCAGCT	CCGTGGATTT	5450
	TCAAATCGCA	ATGGTCCCTG	TTAATGCCAT	TGCTTCAAG	GGACTGTGGG	5500
20	AGAAGACATT	TAAGGCTGAA	GACACGCAAA	CAATACCTTT	CAGAGTGACT	5550
	GAGCAGAAAA	GCAAACTCTG	GCAGATGATG	TACCAGATTG	GTTCATTATA	5600
	AGTGCGATCA	ATGGCTCTTG	AGAAATGAA	GATCTGGAG	CTTCATTGTC	5650
	CCAGTGGAAAC	AATGAGCATG	TTGGTGCTGT	TGCCCTGATG	TGTCTCAGGC	5700
	CTTGAGCAGC	TTGAGAGTAT	AATCAGCTTT	GAATAACTGA	CTGAATGGAC	5750
25	CAGTTCTAGT	ATTATGGAAG	AGAGGAAGGT	CAAAAGTATC	TTACCTCGCA	5800
	TAGAGATGGA	GGAGAAATAC	AACCTCACAT	CTCTCTTAAT	GGCTATGGGA	5850
	ATTACTGACC	TGTTCAAGCT	TTCAAGCAAT	CTGTCTGGCA	TCTCTCACT	5900
	AGGAGCCCTG	AGATATCTC	AAAGCTGTCCA	TGCAGCACAT	GCAGAAATCA	5950
	ATGAAGCGGG	CAGAGATGTG	GTAGGCTCAG	CAGAGGCTGG	AGTGAGATCT	6000
30	ACTGAAGAAAT	TTAGGGCTGA	CCATCCATTTC	CTCTTCTGTG	TCAAGCATAT	6050
	CGAAACCAAC	GCCATTCTCC	TCCTTGGCMG	ATGTGTTTCT	CCGCGGCCAG	6100
	CAGATGACGC	ACCAGCAGAT	GACGCACCCAG	CAGATGACGC	ACCAGCAGAT	6150
	GACGCACCCAG	CAGATGACGC	ACCAGCAGAT	GACGCACCAAC	CATGTATCCT	6200
	GAAAGGCTCT	TGTGGCTGGA	TGCGCTCTCT	GGATGACGAT	GACAAATTTG	6250
35	TGAACCAACA	CCTGTGCGGC	TCACACCTGG	TGGAAGCTCT	CTACCTAGTC	6300
	TGCGGGGAAC	GAGGCTTCTT	CTACACACCC	AAGACCCGCC	GGGAGGCAGA	6350
	GGACTCTGAC	GTGGGGCAGG	TGGAGCTGGG	CGGGGGCCCT	GGTGCAGGCA	6400
	GCCTGCAGCC	CTTGGCCCTG	GAGGGGTCCC	TGCAGAAGCG	TGGCATTTGT	6450
	GAACAATGCT	GTACCAAGAT	CTGCTCCCTC	TACCAGCTGG	AGAACTACTG	6500
40	CAACTAGGGC	GCCTGGATCC	AGATCACTTC	TGGCTAATAA	AAAGATCAGAG	6550
	CTCTAGAGAT	CTGTGTGTTG	GTITTTTGTG	GATCTGCTGT	GCCTTAGGAT	6600
	TGCCAGCCAT	CTGTGTGTTG	CCCTCCCTCC	GTGCCTTCTC	TGACCCTGGA	6650
	AGGTGCCACT	CCCCTGTCC	TTTCTAATA	AAATGAGGAA	ATTGCATCGC	6700
	ATTGTCTGAG	TAGGTGTCAAT	TCTATTCTGG	GGGGTGGGGT	GGGGCAGCAC	6750
45	AGCAAGGGGG	AGGATTGGGA	AGACATAGC	AGGCATGCTG	GGGATGCGGT	6800
	GGGCTCTATG	GGTACCTCTC	TCTCTCTCTC	TCTCTCTCTC	TCTCTCTCTC	6850
	TCTCTCGGTA	CCTCTCTCGA	GGGGGGGGCC	GGTACCCAAAT	TGCGCCTATA	6900
	GTGAGTCGTA	TTACGCGCGC	TCAGTGGCGC	TGCTTTTACA	ACGTGCTGAC	6950
50	TGGGAAGAAC	CTGGCGTTAC	CCAATTTAAT	CGCCTTGACG	CACATCCCCC	7000
	TTTCGCGAGC	TGGCGTAATA	GCGAAGAGGC	CCGCAACGAT	CGCCCTTCCC	7050
	AACAGTTGCG	CAGCGTGAAT	GGCGAATGGA	AATGTGAAGC	GTTAATATTT	7100
	TGTTAAAAAT	CGCGTTAAAT	TTTGTGTAAT	TGAGTCACTT	TTTAAACCAA	7150
	TAGGCCGAAA	TCGGCAAAAT	CCCTTATATA	TCAAAGAAAT	AGACCGAGAT	7200
	AGGTTTGAGT	GTTGTTCCAG	TTTGGAAACA	GAGTCCACTA	TTAAAGAACG	7250
55	TGAGCTCCAA	CGTCAAAGGG	CGAAAAACCG	TCTATCAGGG	CGATGGCCCC	7300
	CTACTCCGGG	ATCATATGAC	AAAGATGTGA	TCCACCTTAA	CTTAATGATT	7350
	TTTACCAAAA	TCATTAGAGG	ATTCACTAGT	GCTCAGGGTC	AACGAGAAAT	7400
	AACATTCGCT	CAGGAAAGCT	TATGATGATG	ATGTGCTTAA	AAACTTACTC	7450
	AATGGCTGGT	TATGATATAT	GCAATACATG	CGAAAAACCT	AAAAAGAGCT	7500
60	GCGGATAAAA	AAGGCCAAT	TATGCTATT	TACCGCGGCT	TTTTATTGAG	7550
	CTTGAAAGAT	AAATAAAATA	GATAGTTT	ATTGAAAGCT	AAATCTTCTT	7600

	TATCGTAAAA	AATGCCCTCT	TGGGTATCA	AGAGGGTCAT	TATATTTGCG	7650
	GAATTAACAT	CATTTGGTGA	CGAAATAACT	AAGCACTTGT	CTCCTGTTTA	7700
	CTCCCCTGAG	CTTGAGGGGT	TAACATGAAG	GTCACTCGATA	GCAGGATAAT	7750
5	AATACAGTAA	AACGCTAAAC	CAATAATCCA	AATCCAGCCA	TCCCAAATGT	7800
	GTAGTGAATG	ATTATAAATA	ACAGCAAAAC	GTAAATGGGCC	AATAACACCG	7850
	GTTCGATTGG	TAAAGCTCAC	CAATAATCCC	TGTAAGAGCAC	CTTGCTGAATG	7900
	ACTCTTTGTT	TGGATAGACA	TCACCTCCCTG	TAATGCAAGT	AAAGCGATCG	7950
	CACCACCAGC	CAATAAAATT	AAAAACAGGA	AAACTAACCA	ACCTTCAGAT	8000
	ATRAACGCTA	AAAAGGCAAA	TGCACTACTA	TCCTCAATAA	ATCCGAGCAG	8050
10	TACTGCGGTT	TTTTCGCCCC	ATTTAGTGGC	TATTCTTCTC	GCCACAAAGG	8100
	CTTGGAATAC	TGAGGTAA	AGACCAAGAC	CGGTAATGA	AAAGCCAACC	8150
	ATCATGCTAT	TCCATCCAAA	ACGATTTTCG	GTAATATAGA	CCCACACCGT	8200
	TGCGGGAATT	TGGCTATCA	ATTGCGCTGA	AAAATAAATA	ATCAACAAAA	8250
	TGGCATCGTT	TTAATAAAG	TGATGTATAC	CGAATTCAGC	TTTTGTTCCT	8300
15	TTTAGTGAGG	GTTAATTGCG	CGCTTGGCGT	AATCATGGTC	ATAGCTGTGT	8350
	CCTGTGTGAA	ATTGTTATCC	GCTCACAAAT	CCACACAACA	TACGAGCCGG	8400
	AAGCATTAAG	TGTAAAGCCT	GGGTCGCTA	ATGAGTGAGC	TAACACACAT	8450
	TAATTGCGTT	CGGCTCACTG	CCGCTTTTCC	AGTCGGGAAA	CCTGTGCTGC	8500
	CAGCTGCATT	AATGAATCGG	CCAACGCGCG	GGGAGAGCGG	GTTTTCGGTAT	8550
20	TGGCGCCTCT	TCCGCTTCCT	CGCTCACTGA	CTCGCTCGCG	TCCGTCGCTT	8600
	GGCTGCGGCG	AGCGGTATCA	GCTCACTCAA	AGGCGGTAAAT	ACGTTTATCC	8650
	ACAGAATCAG	GGGATAACGC	AGGAAGAAGC	ATTGTAGCAA	AAGGCCACGA	8700
	AAAGGCGCAG	AACGTAATA	AGGCGCGGTT	GCTGCGGCTT	TTCATAGGCG	8750
	TCCGCCCCCT	TGACGAGCAT	CACAAAAATC	GACGCTCAAG	TCAGAGGTGG	8800
25	CGAAACCCGA	CAGGACTATA	AAGATACCAG	GGGTTTCCCC	CTGGAAGCTC	8850
	CTCGTGCGC	TCTCCTGTTC	CGACCCCTGC	GCTTACCGGA	TACCTGTCCG	8900
	CCTTTCTCCC	TTGCGGAAGC	GTGGCGCTTT	CTCATAGCTC	ACGCTGTAGG	8950
	TATCTCAGTT	CGGTGTAGGT	CGTTGCGTCC	AAGCTGGGCT	GTGTGCACGA	9000
	ACCCGCCGTT	CAGCCCGACC	GCTGCGCTTT	ATCCGGTAAC	TATCGTCTTG	9050
30	AGTCCAACCC	GGTAAGACAC	GACTTATCGC	CAGTGGCAGC	AGCCACTGGT	9100
	ACAGGATTA	GCAGAGCGAG	GTATGTAGCG	GGTGCTACAG	AGTTCTTGAA	9150
	GTGTGCGCT	AACCTACGGT	ACACTAGAAG	GACAGTATTT	GGTATCTCGG	9200
	CTCTGCTGAA	GCCAGTTACC	TTGCGAAAAA	GAGTTTGGTAG	CTCTTGATCC	9250
35	GGCAACAA	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	9300
	GATTACGCG	AGAAAAAAG	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	9350
	CGGGGTCTGA	CGCTCAGTGG	AACGAAACT	CACGTTAAGG	GAITTTTGTC	9400
	ATGAGATTAT	CAAAAAGGAT	CTTCACTAG	ATCCTTTTAA	ATTAAAAATG	9450
	AAGTTTAA	TCAATCTAAA	GTATATATGA	GTAACATTGG	TCTGACAGTT	9500
	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	9550
40	TATCATATG	TTGCTGAAT	CCCGTCGTG	TAGATAACTA	CGATACGGGA	9600
	GGGCTTACCA	TCTGGCCCCA	GTGCTGCAAT	GATACCCGGA	GACCCACGCT	9650
	ACCGGCTCC	AGATTTATCA	CGAATAAAC	AGCCAGCCGG	AAGGGCCGAG	9700
	CGCAGAAAGT	GCTCTGCAAC	TTTATCCGCC	TCCATCCAGT	CTATTAAATG	9750
	TTGCGGGGAA	GCTAGAGTAA	GATGTTGCGC	AGTTTAAGT	TTGCGCAACG	9800
45	TGTTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGTGATG	9850
	GCTTCATCA	GCTCGGTTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	9900
	CATGTTGTGC	AAAAAAGCGG	TTAGCTCCTT	CGGTCCTCG	ATCGTTGTCA	9950
	GAGTAAAGTT	GGCGCAGTGT	TTATCACTCA	TGGTTATGGC	AGCACTGCAT	10000
50	AATTCTCTTA	CTGTCATGCC	ATCCGTAAGA	TGCTTTTCTG	TGACTGGTGA	10050
	GTATCCAAAC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT	10100
	TTGCGCGCG	GTCAATACGG	GATAATACCG	CGCACATAG	CAGAAGCTTTA	10150
	AAAGTGCTCA	TCAATTGGAAA	ACGTTCTTCG	GGGCGAAAAA	TCTCAAGGAT	10200
	CTTACCGCTG	TTGAGATCCA	TTTCGATGTA	ACCACTCGT	GACCCCAACT	10250
	GATCTTCAGC	ATCTTTTACT	TTTCCACAGC	TTTCTGGGTC	AGCAAAAAACA	10300
55	GGAAGGCAAA	ATGCGGCAAA	AAAGGGAATA	AGGCGCACAC	GGAATGTGTG	10350
	AATAGTCATA	CTCTTCTCTT	TTCAATATTA	TGAAGCATT	TATCAGGGTT	10400
	ATTGTCATCT	GAGCGGATAC	ATATTGAAT	GTATTTAGAA	AAATAACAAA	10450
	ATAGGGGTTT	CGCGCACATT	TCCCGCAAAA	GTGCCAC		10487
60	SEQ ID NO:45 (pTnMod(Oval/ENT tag/P146/PA) - Chicken)					

	CTGACGCGCC	CTGTAGCGGC	GCATTAAGCG	CGCGGGTGT	GGTGGTTACG	50
	CGCAGCGTGA	CCGCTACACT	TGCCAGCGCC	CTAGCGCCCG	CTCCTTTCGC	100
	TTTCTTCCCT	TCCTTTCTCG	CCACGTTCCG	CGGCATCAGA	TTGGCTATTG	150
5	GCCATTGCAT	ACGTGTATCA	CATATCATAA	TATGTACATT	TATATTGGCT	200
	CATGTCCAAC	ATTACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	250
	TAGTATCAAA	TTACGGGGTC	ATTAGTTTCA	AGCCCATATA	TGGAGTTCCG	300
	CGTTACATAA	CTTACGGTAA	ATTGGCCGCG	TGGCTGACCG	CCCAACGACC	350
	CCCGCCCAT	GACGTCAATA	ATGAGCTATG	TTCCTCATGT	AACGCCAATA	400
	GGGACTTTCC	ATTGACGTCA	ATTGGGTGGAG	TATTTACGCT	AAACTGCCCA	450
10	CTTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTAGCGCC	CCATTGTGAC	500
	TAAATGACGG	TAAATGGGCC	GCCCTGGCATT	ATGCCCAGTA	CATGACGATA	550
	TGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC	600
	GATGGTGTATG	CGGTTTGGC	AGTACATCAA	TGGCGGTGGA	TAGCGGTTTG	650
	ACTCA CGGGG	ATTTCCCAAGT	CTOCACCCCA	TTGACGTCAA	TGGGAGTTTG	700
15	TTTTTGCACCC	AAAATCAACG	GGACTTTCCA	AAATGTGCTA	ACAACCTCCG	750
	CCCATTGACG	CAAAATGGGG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	800
	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CTCGGAGACG	CCATCCACGC	850
	TGTTTGTGACC	TCATAGAAAG	ACACCGGAGC	CGATCCAGCC	TCGCGGGCGG	900
	GGAAACGGTGC	ATTGGAACGC	GGATTCCCGC	TGCCAAGAGT	GACGTAACTA	950
20	CCGCCATATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA	1000
	CTGTTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTTCTCTAT	GCTATAGGTT	1050
	ATGGTAATAGC	TTAGCCTATA	GGTGTGGGTT	ATTGACCATT	ATTGACCATT	1100
	CCCTTATTGG	TGACGATACT	TTCCATTACT	AATCCATAAC	ATGGCTCTTT	1150
	GCCACAACATA	TCTCTATTGG	CTATATGCCA	ATACTCTGTC	CTTCAGAGAC	1200
25	TGACACGGAC	TCTGTATTTT	TACAGGATGG	GGTCCCATTT	ATTATTATCA	1250
	AATTACATAA	TACAACAACG	CGTCCCGCG	TGCCCGCAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAATCTCCG	GTACGTGTTC	CGGACATGGG	1350
	CTCTCTCCCG	TGAGCGCGCG	AGCTTCCACA	TCCGAGCCCT	GGTCCCATGC	1400
	CTCCAGCGGC	TCAATGTCGC	TGGGCAGCTC	CTTGCTCTTA	ACAGTGGAGG	1450
30	CCAGAGCTGC	GCACGACACA	ATGCCCAACA	CCACCAAGTT	CGCGCAACAG	1500
	GCGCTGGGCG	TAGGGTATGT	GTCTGAAAT	GAGCGTGGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGAAA	GACTTAAGGC	AGCGGCAGAA	TAGGATGCAG	1600
	GCAGCTGAGT	TGTTGTATTC	TGATAAGAGT	CAGAGGTAAAC	TCCCGTTGCG	1650
	GTGCTGTAA	CGGTGGAGGG	CAGTGTAGTC	TGAGCAGTAC	TCGTTGCTGC	1700
35	CGCGCGCGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCCCTT	1750
	CCATGGGTCT	TTTCTGCAGT	CACCGTCGGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGATT	CTCTTTACCA	ATTCGCGCC	GAATTACACT	TAAAACGACT	1850
	CAACAGCTTA	AGTTTGGCTT	GCCACGCATT	ACTTGACTGT	AAAACCTCTA	1900
	CTCTTACCGA	ACTTGGCGGT	AACCTGCCAA	CCAAAGCGAG	AACAAAACAT	1950
40	AACATCAAA	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAAGA	2000
	GCAGCTCGCT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TCGGGAATAC	2050
	GATGCCCAT	GTACTGTGTG	ACTGGTCTGA	TATTCGTGAG	CAAAAACGAC	2100
	TTATGCTATT	GGGAGCTTCA	GTGCACTAC	ACGCTCGTTC	TGTTACTCTT	2150
	TATGAGAAAG	CGTTCCCGCT	TTACAGACAA	TGTTCAAAGA	AAGCTCATGA	2200
45	CCAAATTTCTA	GCCGACCTTG	CGAGCAITCT	ACCGAGTAAC	ACCACACCGC	2250
	TCATTGTGAG	TGATGCTGGC	TTTAAAGTGC	CATGGTATAA	ATCCGTTGAG	2300
	AAGCTGGGTT	GGTACTGGTT	AAGTCGAGTA	AGAGGAAAGG	TACAATATGC	2350
	AGACCTTAGGA	GCGGAAAACT	GGAAACCTAT	GACCAACTTA	CATGATATGT	2400
	CATCTAGTCA	CTCAAGAGCT	TTAGGCTATA	AGAGGCTGAC	TAAAAGCAAT	2450
50	CCRAATCTCAT	GCCAAATTTCT	ATTGTATAAA	TCTGCTCTTA	AAGGCCGAAA	2500
	AAATCAGCGC	TGACACGCGA	CTCATTTGTA	CCACCCGTCA	CCTAAATATCT	2550
	ACTCAGCGTC	GGCAAGGAGG	CCATGGGTTT	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTCGAA	CACCCAAACA	ACTTGTATAA	ATCTATTGCA	AGCGAATGCA	2650
	GATTGAAGAA	ACCTTCCGAG	ACTTGAAGAG	TCTTGCCTAC	GGACTAGGCC	2700
55	TACGCCATAG	CGGAACGAGC	AGCTCAGAGC	GTTTTGATAT	CATGCTGCTA	2750
	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	TTGCGGGGCG	TTTCATGCTCA	2800
	GAACAAAGGT	TGGGACAAAG	ACTTCCAGCG	TAAACAGTGC	AGAAATCGAA	2850
	ACGTACTCTC	AACAGTTCCG	TTAGGCATGG	AAGTTTTCGC	GCATTCTGCG	2900
	TACACAATAA	CAAGGGGAAGA	CTTACTCGTG	GCTGCAACCC	TACTAGTCTA	2950
60	AAATTTATTC	ACACATGGTT	ACGCTTGGG	GAAATATGTA	TAATGATCCA	3000
	GCTACCTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTGTGTTGG	3050

TTTTGTGG ATCTGCTGTG CCTTCTAGTT GCCAGCCATC TGTGTTTGC 3100
 CCTCCCTGGG TGCCCTTCTT GACCCCTGGAA GGTGCCACTC CCATGTCTCT 3150
 TTCCATAATA AATGAGGAAA TTGCATCGCA TTGTCGTAGT AGGTGTCTATT 3200
 CTATCTCTGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATGGGAAA 3250
 5 GACAAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATTG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAGAAA 3400
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAGAAA 3400
 TTGACCCGGT GACCAAAGGT GCCTTTTATC ATCACTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCCGTGTA TAAGCAGCAA TTAATTATGA TTGATGCCCTA 3500
 10 CATCACAA CA AAACTGATT TAACAAATGG TTGGTCTGCC TTGAAAAGTA 3550
 TATTTGAACA TTATCTTGAT TATATTATTG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTTGGT TGGAAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGTAA ACGCCATTGT CAGCAAAATTG 3700
 ATCCAGAGGA ACCAACTTAA AGCTTTCTCTG ACGGAATGTT AATTCTCGTT 3750
 15 GACCCCTGAGC ACTGATGAAT CCCCTAATGA TTTTGGTAAA AATCAITTAG 3800
 TTAAGCTGGA TACACATCTT GTCATATGAT CCCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCOAGG CTTTACACTI TATGCTTCGG GCTCGTATGT 3900
 TGTGTGGAAT TGTAGCGGGA TAACAATTTT ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGGCG AATTAACCTC CACTAAAGGG AACAAAGCT 4000
 20 GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTAGTGGG TCCCCCGGG 4050
 AGGTCAAGAT GGTTCCTTTA CTGTTTGICA ATTCTATTAT TTCAATACAG 4100
 AACCAATAGT TCTATAACTG AAATATATTI GCTATTGTAT ATTATGATTG 4150
 TCCCTCGAAG CATGAACACT CCTCCAGCTG AATTTACAAA TTCTCTGTCT 4200
 ATCTGCCAGG CCATTAAGTT AITCATGGAA GATCTTTGAG GAACACTGCA 4250
 25 AGTTCATATC ATAAACACAT TTGAAATGTA GTATTGTTTT GCATTGTATG 4300
 GAGCTCATGT TTGCTGTATC CTCAGAAAAA AAGTTTGTTA TAAAGCATTC 4350
 ACACCCATAA AAAGATGAT TTAATAATTC CAGCTATAGG AAAGAAAGTG 4400
 CGCTCTCTCT TCACTCTAGT CTCAGTTGGC TCCTTCACAT GCATGCTTCT 4450
 TTAATTCTCC TATTTTGTGA AGAAATAAT AGGTCACGTC TTGTTCTCAC 4500
 30 TTAATGCTCG CTAGCATGCG CTCAGATGCA CGTTGTAGAT ACAAGAAGCA 4550
 TCAAAATGAA CAGACTCTTG GTCTGTTACT ACAACCATAG TAATACAGCA 4600
 ACTAAGTAAT AATGCTAAT TATGTTTCCC ATCTCTAAGG TTCCACATT 4650
 TTTCTGTTT CTTAAAGATC CCATTATCTG GTTGTAACTG AAGCTCAATG 4700
 35 GAACATGAGC AATATTCCC AGTCTTCTCT CCCATCCAAC AGTCTCTGAT 4750
 GATTAGCAGA ACAGGCAGAA AACACATTGT TACCCAGAAT TAAAACTTAA 4800
 TATTTGCTCT CCATTCAATC CAAAATGGAC CTATTGAAAC TAAATCTTAA 4850
 CCCAATCCCA TTAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG 4900
 GGAACCTGTG GGTGGGTGAC AATTCAGGCT ATATATTCOC CAGGGCTCAG 4950
 40 CGGATCCATG GGCTCCATCG GCGCAGCAAG CATGGAATTG TGTGTTGATG 5000
 TATTCAAGGA GCTCAAAGTC CACCATGCCA ATGAGAACAT CTCTACTGCG 5050
 CCCATTGCCA TCAATGTCAG TCTAGCCATT GTATACCTGG GTGCAAAAGA 5100
 CAGCACCAGG ACACAGATAA ATAAGGTTGT TCGCTTTGAT AAACCTCCAG 5150
 GATTCCGAGA CAGTATTGAA GCTCAGTGTG GCACATCTGT AAAGCTTCAC 5200
 45 TCTTCACTTA GAGACATCTT CAACCAATCT ACCAAACCAA ATGATGTTTA 5250
 TTGCTCAGC CTTGCCAGTA GACTTTATGC TGAAGAGAGA TACCCAATCC 5300
 TGCCAGAATA CTTGCAGTGT GTGAAGGAAC GCCAGAGAGC TCACTAATTC 5400
 CCTATCAACT TTCAAACAGC TGCAAGATCAA CAGAAATGTC CTTCAGCCAA 5450
 50 GCTCCGTGGA TTCTCAAATC GCAATGGTTC TGGTTAATGC CATGTCTTTC 5500
 AAGGCACTGT GGGAGAAAAC ATTTAAGGAT GAAGACACAC AAGCAATGCC 5550
 TTTCAAGATG ACTGAGCAAG AAAGCAAACC TGTGCAGATG ATGTACCAGA 5600
 TTGTTTATT TAGATGGCA TCAATGGCTT CTGAGAAAT GAAGATCTGT 5650
 GAGCTTCCAT TTGCCAGTGG GACAATGAGC ATGTGTGTCG TGTGCTGTA 5700
 55 TGAAGTCTCA GGCTTTGAGC AGCTTGAGAG TATAATCAAC TTTGAAAAAC 5750
 TGACTGAATG GACCAATTCT AATGTTATGG AAGAGAGGAA GATCAAAGTG 5800
 TACTTACCTC GCATGAAGAT GGAGGAAAAA TACAACCTCA CATCTGCTTT 5850
 AATGGCTATG GGCATTACTG ACGTGTTTAG CTCTTCAGCC AATCTGTCTG 5900
 GCATCTCTCT AGCAGAGAGC CTGAAGATAT CTCAGGCTGT CCATGCAGCA 5950
 CATGCAGAAA TCAATGAAGC AGGCAGAGAG GTGGTAGGGT CAGCAGAGCG 6000
 60 TGGAGTGGAT GCTGCAAGCG TCTCTGAAGA ATTTAGGGCT GACCATCCAT 6050
 TCCTCTCTG TATCAAGCAC ATCGCAACCA ACGCGTCTCT CTTCTTTGGC 6100

	AGATGTGTTT	CCCCTCGCG	GCCAGCAGAT	GACGCACCAG	CAGATGACGC	6150
	ACCAGCAGAT	GACGCACCAG	CAGATGACGC	ACCAGCAGAT	GACGCACCAG	6200
	CAGATGACGC	AACACATGT	ATCCTGAAAG	GCTCTGTGG	CTGGATCGGC	6250
	CTGCTGGATG	ACGATGACAA	AAAATACAAA	AAAGCACTGA	AAAAACTGGC	6300
5	AAAACGTGCT	TAATGAGGGC	GCCTGGATCC	AGATCACTTC	TGGCTAATAA	6350
	AAGATCAGAG	CTCTAGAGAT	CTGTGTGTGT	GTTTTGTGTG	GATCTGCTGT	6400
	GCCTCTTAGT	TGCCAGOCAT	CTGTGTGTGT	CCCCCTCCCC	GTGCCCTTCCT	6450
	TGACCTCGGA	AGGTGCCACT	CCCAGTGTCC	TTTCTTAATA	AAATGAGGAA	6500
	ATTGTACATCG	ATTGTCTGAG	TAGGTGTCTAT	TCTATTCTCG	GGGGTGGGGT	6550
10	GGGGCAGCAC	AGCAAGGGGG	AGGATTTGGGA	AGACAATAGC	AGGCATGCTG	6600
	GGGATCGCGT	GGGCTCTATG	GGTACCTCTC	TCTCTCTCTC	TCTCTCTCTC	6650
	TCTCTCTCTC	TCTCTCGGTA	CCTCTCTCGA	GGGGGGGGCC	GGTACCCAAAT	6700
	TGCCGCTATA	GTGAGTCGTA	TTACGGCGCG	TCAGTGGCGC	TGCTTTTACA	6750
	ACGTCGTGAC	TGGGAAAAAC	CTGGCGTTAC	CCAACCTTAAT	CGCCTTGCAG	6800
15	CACATCCCCC	TTTCGCCAGC	TGSCGTAAAT	GCGAAGAGGC	CCGACCCGAT	6850
	CGCCCTTCCC	AACAGTTGCG	CAGCCTGAAT	GGCGAATGGA	AATTGTAAAG	6900
	GTTAATATT	TGTTAAATTT	CGCGTTAAAT	TTTGTGTAAA	TCAGCTCAT	6950
	TTTAAACCAA	TAGGCCGAAA	TGCGCAAAAT	CCCTTATAAA	TCAAAGAAAT	7000
	AGACCAGAGAT	AGGGTTGAGT	GTGTGTCCAG	TTTGAACCAA	GAGTCCACTA	7050
20	TTAAAGAACG	TGGAATCCAA	CGTCAAGAGG	CGAAAAACCG	TCTATCAGGG	7100
	CGATGGCCCA	CTATCCGGG	ATCATATGAC	AAGATGTGTA	TCCACCTTAA	7150
	CTTAATGATT	TTTACCAAAA	TCATTAGGGG	ATTTCATGAT	GCTCAGGGTC	7200
	AACGAGAAAT	AACATTCGGT	CAGGAAGAGT	TATGATGATG	ATGTGCTTAA	7250
	AAACTTACTC	AATGGCTGGT	TATGCATATC	GCAATACATG	CGAAAAACCT	7300
25	AAAAGAGCTT	CGCGATAAAA	AAGGCCAATT	TATTGCTATT	TACCGCGGCT	7350
	TTTATTGAG	CTTGAAGAT	AAATAAAAT	GATAGGTTT	ATTTGAAGCT	7400
	AAATCTTCTT	TATCGTAAAA	AATGCCCTCT	TGGGTTATCA	AGAGGGTCTAT	7450
	TATATTTCG	GGAATAACAT	CATTTGGTGA	CGAAATAACT	AGCACTTGT	7500
	CTCCTGTTTA	CTCCCTGAG	CTTAGGGGCT	TAACATGAAG	GTACATCGATA	7550
30	CGAGGATAAT	AATACAGTAA	AACGCTAAAC	CAATAATCCA	ATCCACGACA	7600
	TCCCAAATTG	GTAGTGAATG	ATTATAAATA	ACACCAACAA	GTAATGGGCC	7650
	AATAACACCG	GTTGCATTGG	TAAGGCTCAC	CAATAATCCC	TGTAAGGAC	7700
	CTTGCTGATG	ACTCTTTGTT	TGGATAGACA	TCACCTCCGT	TAATGAGGTT	7750
35	AAAGCGATCC	CACCACCAGC	CAATAAAATT	AAAAACAGGA	AAACTAACCA	7800
	ACCTTCAGAT	ATAAACGCTA	AAAAGGCAAA	TGCACTACTA	TCTGCAATAA	7850
	ATCCGAGCAG	TACTGCCGTT	TTTTGCCCC	ATTATGTGGC	TATTCTTCCT	7900
	GCCACAAAGG	CTTGAATATC	TGAGTGTAAA	AGACCAAGAC	CCGCTAATGA	7950
	AAAGCCAACC	ATCATGCTAT	TCCATCCAAA	ACGATTTTCG	GTAATATAGA	8000
	CCACACCGT	TGCCGGAATT	TGSCCTATCA	ATTGCGCTGA	AAAAATAATA	8050
40	ATCAACAAA	TGGCATCGTT	TTAAATAAAG	TGATGTATAC	CGAATTCAGC	8100
	TTTTGTTCCT	TTTAGTGAGG	GTTAATTGCG	CGCTTGGCGT	AATCATGGCT	8150
	ATAGCTGTTT	CCTGTGTGAA	ATTGTTATCC	GCTCACAATT	CCACCAACAA	8200
	TACGAGCCGG	AAGCATAAAG	TGTAAGGCTT	GGGTGCGCTA	ATGAGTGAGC	8250
	TAATCTACAT	TAATTGCGTT	CGCCTCACTG	CCCGCTTTCC	AGTGGGGAAA	8300
45	CCTGTCTGTC	CAGCTGCATT	AATGAATCGG	CCACGCGCGC	GGGAGAGGCG	8350
	GTTTGCCTAT	TGGCGCTCT	TCCGCTTCCT	CGCTCACTGA	CTCGCTCGCG	8400
	TGCGTCTGTC	GGCTGCGCGC	AGCGGTATCA	GCTCACTCAA	AGGCGGTAAAT	8450
	ACGGTTTATC	ACAGAACTAG	GGGATAACGC	AGGAAAGAAC	ATGTGAGCAA	8500
	AAAGGCCAGCA	AAAGGCCAGG	AACCGTAAAA	AGGCGCGGTT	GCTGGCGTTT	8550
50	TTTCCATAGG	TCCGCCCCCC	TGACGAGCAT	CACAAAAATC	GACGCTCAAG	8600
	TCGAGGTGGG	CGAAACCCGA	CAGGACTATA	AAGATACCAG	CGGTTTCCCC	8650
	CTGGAAGCTC	CCTCGTGCGC	TCTCTGTGTC	CGACCTGCGC	GCTTACCGGA	8700
	TACCTGTCCG	CCTTTCTCCC	TTCCGGGAAG	GTGGCGCTTT	CTCATAGCTC	8750
	ACGCTGTAGG	TATCTCAGTT	CGGTGTAGGT	CGTTTCGCTC	AAGCTGGGCT	8800
55	GTGTGCACGA	ACCCCGCGTT	CAGCCCGAGC	CGTCGCGCTT	ATCCGGTAAC	8850
	TATCTGCTTG	AGTCCAAACC	GGTAAAGAC	GACTTATGCG	CAGTGGCAGC	8900
	AGCCACTGGT	AACAGGATTA	CGAGAGCGAG	GTATGTAGGC	GGTGTACAGC	8950
	AGTTCTTGAA	GTGTGGGCTC	AACATACGGT	ACACTAGGAG	GACAGTATT	9000
	GGTATCTGCG	CTCTGCTGAA	GCCAGTTACC	TTGCGAAAAA	GAGTTGTGAG	9050
60	CTCTTGTATCC	GGCAACAAA	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	9100
	GCAAGCAGCA	GATTACGCGC	AGAAAAAAG	GATCTCAAGA	AGATCTTTTG	9150

	ATCTTTTCTA	CGGGGCTGA	CGCTCAGTGG	AACGAAAAC	CACGTTAAGG	9200
	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACTAG	ATCCTTTTAA	9250
	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTTAACTTGG	9300
	TCGTACAGTT	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	9350
5	TCTATTTTCT	TCAATCCATG	TTGCCTGACT	CCCGTGTGTC	TAGATAACTA	9400
	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA	GTGCTGCAAT	GATACCGGCA	9450
	GACCCAGCCT	CACCGCTCC	AGATTATACA	GCAATAAAC	AGCCAGCCGG	9500
	AAGGCCGAG	CGCAGAAGTG	GTCTGCAAC	TTTATCCGCC	TCCATCCAGT	9550
	CTATTATATG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTCCGC	AGTTAATAGT	9600
10	TTGCCCAACG	TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCTGC	9650
	GTTTGGTATG	GCTTCATICA	GCTCCGGTTC	CCAACGATCA	GGCGGAGTTA	9700
	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG	TTAGTCTCCT	CGGTCTCTCG	9750
	ATCGTGTGTA	GAAATAAGTT	GGCGCAGTGG	TTATCACTCA	TGGTTATGCG	9800
	AGCACTGCAT	AATTCTCTTA	CTGTCAATGC	ATCCGTAAGA	TGCTTTTCTG	9850
15	TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATCGGCGGA	9900
	CCGAGTTGCT	CTTGCCCGCG	GTCAATACGG	GATAATACCG	CGCCACATAG	9950
	CAGAAGTTTA	AAAGTGCTCA	TCAATTGGAA	ACGTTCTTCG	GGGCGAAAC	10000
	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA	GTTCGATGTA	ACCCACTCTG	10050
	GCACCAAGCT	GATCTTCAGC	ATCTTTTACT	TTCAACGCGG	TTTCTGGGTG	10100
20	AGCAAAAACA	GGAAAGCAAA	ATGCCGCAAA	AAAGGGAATA	AGGCGCACAC	10150
	GGAAATGTTG	AATACTCATA	CTCTTCTCTT	TTCAATATTA	TTGAAGCAAT	10200
	TATCAGGGTT	ATTGTCTCAT	GAGCGGATAC	ATATTTTGAAT	GTATTTTAGAA	10250
	AAATAAACAA	ATAGGGGTTT	CGCGCACATT	TCCCCGAAAA	GTGCCAC	10297

25

SEQ ID NO:46 (pTnMod(Oval/ENT tag/P146/PA) - QUAIL)

	CTGACGCGCC	CTGTAGCGGC	GCAITTAAGCG	CGGCGGGTGT	GGTGGTTACG	50
	CGCAGCGTGA	CCGCTACACT	TGCCAGCGCC	CTAGCGCCCG	CTCCCTTTCCG	100
30	TTTCTTCCCT	TGCTTTCTCG	CCACGTTCCG	CGGCATCAGA	TTGGCTATTG	150
	GCCATTGTCAT	ACGTTGTATC	CATATCATAA	TATGTACATT	TATATTGGCT	200
	CATTGCTCAAC	ATTACCCCCA	TGTGTGACAT	GATTATTGAC	TAGTTATTAA	250
	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT	AGGCCATATA	TGGAGTTCCG	300
	CGTTACATAA	CTTAGCGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	350
35	CCCGCCATT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	400
	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGCT	AAACTGCCCA	450
	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTAACGCC	CCTATTGAGC	500
	TCAATGACGG	TAAATGGCCC	GCTTGGCAIT	ATGCCCAGTA	CATGACCTTA	550
	TGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TGCTATTATC	600
40	CATGTTGATG	CGGTTTGGC	AGTACATCAA	TGGGCGTGGA	TAGCGGTTTG	650
	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG	700
	TTTTTGGACC	AAAATCAACG	GGACTTTCCA	AAATGTGCTA	ACAACCTCCG	750
	CCCATTAGCG	CAAAATGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	800
	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACGC	850
45	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGCGAC	900
	GGAAACGGTC	ATTGGAAACG	GGATTCCCGG	TGCCAAGAGT	GACGTAGTGA	950
	CCGCCCTATG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA	1000
	CGTTTITGG	CTTGGGCGCT	ATACACCCCC	GCTTCTCTAT	GCTATAGGTT	1050
	ATGGTATAGC	TTAGCCCTTA	GGTGTGGGTT	ATTGACCATT	ATTGACCACT	1100
50	CCCTATTGGT	TGACGATACT	TTCCATTACT	AATCCATAAC	ATGGCTCTTT	1150
	GCCACAACATA	TCTCTATTGG	CTATATGCCA	ATACTCTGTC	CTTCAGAGAC	1200
	TGACAGCGAC	TCTGTATTIT	TACAGGATGG	GGTCCCAATT	ATTATTATACA	1250
	AATTACATATA	TACAACAACG	CGGTCCCCCG	TGCCCGCAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAATCTCGG	GTACGTTTTC	CGACATGGGG	1350
55	CTCTTCTCGG	GTAGCGGCGG	AGCTTCCACA	TCCGAGCCCT	GGTCCCATGC	1400
	CTCCAGCGGC	TCATGGTTCG	TGGGCGAGTC	CTTGCTCTTA	ACAGTGGAGG	1450
	CCAGACTTAG	GCACAGACACA	ATGCCACCCA	CCACAGTGTG	GCGCCACAAG	1500
	GCGCGTGGCG	TAGGGTATGT	GTCTGAAATC	GAGCGTGGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGGAA	GACTTAAGGC	AGCGGCAGAA	GAAAGATGCAG	1600
60	GCAGCTGAGT	TGTTGTATTCT	TGATAAGAGT	CAGAGGTAAAC	TCCCGTCTCG	1650
	GTGCTGTTAA	CGGTGGAGGG	CAGTGTAGTC	TGAGCAGTAC	TGCTGTGAGC	1700

	CGCGCGCGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCTTTT	1750
	CCATGGGTCT	TTTCTGAGT	CACCGTCGGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGATT	CTCTTTACCA	ATCTGCGCCC	GAATTACACT	TAAAACGACT	1850
5	CACAGCTTA	AOGTGGCTT	GCCAGCGATT	ACTTGACTGT	AAAACTCTCA	1900
	CTCTTACCGA	ACTTGGCCGT	AACCTGCCAA	CCAAAGCGAG	AACAAAACAT	1950
	AACATCAAAC	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAGA	2000
	GGCAGTCGCT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TCGGGAATAC	2050
	GATGGCCATT	GTACTTGTG	ACTGGCTCGA	TATTCGTGAG	CAAAAAACGAC	2100
10	TTATGGTATT	GGGAGCTTCA	GTGCGACTAC	ACGGTCGTTC	TGTTACTCTT	2150
	TATGAGAAAG	CGTTCGCCCT	TTACAGACAA	TGTTCAAAGA	AAGCTCATGA	2200
	CCAAITTTCTA	GCCGACCTTG	CGAGCATTCT	ACCGAGTAAC	ACCACACCGC	2250
	TCATTGTCTAG	TGATGCTGGC	TTTAAAGTGC	CATGGTATAA	ATCCGTTTGA	2300
	AAGCTGGGTT	GGTACTGGTT	AAGTCGAGTA	AGAGGAAAAG	TACAATATGC	2350
15	AGACCTAGGA	GCGGAAAAC	GGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
	CACTTAGTCA	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAAGCAAT	2450
	CCAATCTCAT	GCCAAATTCT	ATTGTATATA	TCTGCTCTTA	AAGGCCGAAA	2500
	AAATCAGCGC	TCGACACGGA	CTCATTTGTA	CCACCCGTCA	CCTAAAATCT	2550
	ACTCAGCGTC	GGCAAGAGAG	CCATGGGTTT	TAGCAACTAA	CTTACCTGTT	2600
20	GAAITTGAAA	CACCCAAAACA	ACTTGTTAAT	ATCTATTGCA	AGCGAATGCA	2650
	GATTGAGAAA	ACCTTCCGAG	ACTTGAAAG	TCTTGCTTAC	GGACTAGGCC	2700
	TACGCCATAG	CCGAAACGAGC	AGCTCAGAGC	GTTTGTGATAT	CATGCTGTCTA	2750
	ATCGCCCTGA	TGCTTCAACT	AACATGTTTG	CTTGGCGGCG	TTTATGCTCA	2800
	GAAACAAGTT	TGGGACAAGC	ACTTCCAGGC	TAAACAGTCT	AGAAATCTGA	2850
	ACGTACTCTC	AACAGTTGCG	TTAGGCATGG	AAGTTTTCGG	GCATTTCTGC	2900
25	TACAACAATA	CAAGGGAAGA	CTTACTCGTG	GCTGCAACCC	TACTAGCTCA	2950
	AAATTTATTC	ACACATGGTT	ACGCTTTGGG	GAAATATGTA	TAATGATCCA	3000
	GATTTCTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTTGTTTGG	3050
	TTCTTTTGGG	ATCTGCTGTG	CCTTCTAGTT	GCCAGCCATC	TGTTGTTTGC	3100
	CCCTCCCGCC	TGCTTCTCTT	GACCTCGGAA	GCTGCCATCT	CCACTGCTCT	3150
30	TTCTTAATAA	AATGAGGAAA	TTGCATGCGA	TTGTCTGAGT	AGGTGTCATT	3200
	GATTTCTGGG	GGGTGGGGTG	GGGCGACACA	GCAAGGGGGA	GGATTGGGAA	3250
	GACAATAGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAGG	TGCTGAAGAA	3400
35	TTGACCCGGT	GACCAAAGGT	GCCTTTTATC	ATCACTTTAA	AAATAAAAAA	3450
	CAATTACTCA	GTGCGTGTTA	TAAGCAGCAA	TTAATTATGA	TTGATGCCCTA	3500
	CATCACAACA	AAAACGTGAT	TAACAAATGG	TTGGTCTGCC	TTAGAAAGTA	3550
	TATTTGAACA	TTATCTTGAT	TATATTATTG	ATAATTAATA	AAACCTTATC	3600
40	CCTATCCAAAG	AAGTGATGCC	TATCATTTGT	TGGAATGAAC	TTGAAAAAAA	3650
	TTAGGCTTGA	ATACATTACT	GGTAAGGTAA	ACGCCAATGT	CAGCAAAATTG	3700
	ATCCAAGAGA	ACCAACTTAA	AGCTTTCTCG	ACGGAATGTT	AATTTCTCGTT	3750
	GACCTCTAGC	ACTGATGAAT	CCCTTAATGA	TTTTGGTAAA	AATCAATTAG	3800
	TAAAGGTGGA	TACACATCTT	GTCAATATAT	CCCGCTAATG	TGAGTTTACT	3850
45	CACCTCATAG	GCACCCAGG	CTTTACACTT	TATGCTTCCG	GCTCGTATGT	3900
	TGTGTGGAAAT	TGTGAGCGGA	TACAAATTC	ACACAGGAAA	CAGCTATGAC	3950
	CATGATTACG	CAAAGCGCGC	AATTAACCTT	CACATAAGGG	AACAAAAGCT	4000
	GGAGCTTCCAC	CGCGGTGGCG	GCGGCTCTAG	AACATGTGGA	TCCCGCGGGG	4050
	AGGTCAGAAAT	GGTTTCTTTA	CTGTTTGTCA	ATTCTATTAT	TTCAATACAG	4100
	AACAAAAGCT	TCTATAACTG	AAATATATTT	GCTATTGTAT	ATTATGATTG	4150
50	TCCTCGCAAC	CATGAACACT	CCTCCAGCTC	AATTTACAAA	TTCTCTTGTC	4200
	ATCTGCCAGG	CTGGAAGATC	ATGGAAGATC	TCTGAGGAAC	ATTGCAAGTT	4250
	CATCACATAA	ACTCATTTGG	AATTGAGTAT	TAITTTTGCTT	TGAATGGAGC	4300
	TATGTTTTCG	AGTTCCCTCA	GAGAAAAAGC	TGTTTATATA	GCGTCTACAC	4350
	CCATCAAAAG	ATATATTTAA	ATATTCCAAC	TACAGAAAAG	TTTTGCTCGC	4400
55	TCATTCACTCT	GATCTCAGTT	GGTTTCTTCA	CGTACATGCT	TCTTTATTTG	4450
	CTATTATTGT	CAAGAAAATA	ATAGGTCAGG	TCCTGTTCTC	ACTTATCTCC	4500
	TGCTTAGCAT	GCCTTAGATG	CACGTTGTAC	ATTCAAGAA	GATCAATATG	4550
	AACAGACTTC	TGGTCTGTTA	CAACAACACT	AGTAATAAAC	AGACTAACTA	4600
	ATAAATTGCTA	ATTATGTTTT	CCATCTCTAA	GGTTCCACCA	TTTTTCTGTT	4650
60	TAAAGATCCC	ATTATCTGTT	TGTAAGTACG	GCTCAATGGA	ACATGAACAG	4700
	TATTTCTCAG	CTTTTCTCC	AGCAATCCTG	ACGGATTAGA	AGAACTGGCA	4750

	GAAAAACACTT	TGTTACCCAG	AATTAAAAAC	TAATATTGTC	TCGCCCTTCA	4800
	ATCCAAAATG	GACCTATTGA	AACTAAAAAT	TGACCCAATC	CCATTAAATT	4850
	ATTTCTATGG	CGTCAAAGGT	CAAACTTTTG	AAGGGAACCT	GTGGGTGGGT	4900
	CCCAATTTCAG	CGTATATATT	CCCCAGGGCT	CAGCCAGTGG	ATCCATGGGC	4950
5	TCCATCGGTG	CAGCAAGCAT	GGAAATTTGT	TTTGATGTAT	TCAAGGAGCT	5000
	CAAACTCCAC	CATGCCAATG	ACAACATGCT	CTACTCCCCC	TTTGCCATCT	5050
	TGTCAACTCT	GGCCATGGTC	TTCCTAGGTG	CAAAAGACAG	CACCAAGGAC	5100
	CAGATAAATA	AGGTGTGTCA	CTTTGATAAA	CTTCAGGAT	TCCGAGACAG	5150
	TATTGAAGCT	CAGTGTGGCA	CATCTGTAAA	TGTTCACTCT	TCACTTAGAG	5200
10	ACATACTCAA	CCAAATCACC	AAACAAAATG	ATGCTTATTC	GTTCAAGCTT	5250
	GCCCATGAGC	TTTATGCTCA	AGAGACATAC	ACAGTCGTGC	CGGAATATCT	5300
	GCATGTGTGT	AAGGAACGTG	ATAGAGGAGG	CTTAGAATCC	GTCAACTTTC	5350
	AAACAGCTGC	AGATCAAGCC	AGAGGCCCTC	TCAATGCTCG	GGTAGAAAGT	5400
	CAGACAAAAG	GAATTATCAG	AAACATCCTT	CAGCCAAGCT	CCGTGTGATT	5450
15	TCAAATGCGA	ATGCTCCTGG	TTAATGCCAT	TGCTTTCAGG	GGACTGTGGG	5500
	AGAAAGCATT	TAAGGCTGAA	GACACGCAAA	CAATACCTTT	CAGAGTGACT	5550
	GAGCAGAGAA	GCAAACTGCT	GCAGATGATG	TACCAAGTTG	GTTCAATTTA	5600
	AGTGGCATCA	ATGGCTTCTG	AGAAAAATGA	GATCCTGGAG	CTTCCATTGG	5650
	CCAGTGGAAAC	AATGAGCATG	TTGGTGCTGT	TGCTGTGATG	TGCTCTCAGC	5700
20	CTTGAGCAGC	TTGAGAGTAT	AATCAGCTTT	GAAAACTGTA	CTGAATGGAC	5750
	CAGTTCTAGT	ATTATGGAA	AGAGGAAGGT	CAAGGTGATC	TTACCTCGCA	5800
	TGAAGTAGGA	GGAGAAATAC	AACCTCACAT	CTCTCTTAAT	GGCTATGGGA	5850
	CTTACTGACC	TGTTTCAGCT	TTACGCCAAT	CTGTCTGGCA	TCTCCTCACT	5900
	AGGGAGCCTG	AAGATATCTC	AAGCTGTCCA	TGCAGCACAT	GCAGAAATCA	5950
25	ATGAAGCGGG	CAGAGATGTG	GTAGGCTCAG	CAGAGGCTGG	AGTGGATGCT	6000
	ACTGAAGAAT	TTAGGGCTGA	CCATCCATTC	CTCTTCTGTG	TCAAGCACAT	6050
	CGAAACCAAC	GCCATTCTCC	TCTTTGGCAG	ATGTTTCTCT	CCGGGGCCAG	6100
	CAGATGACCG	ACCAGCAGAT	GACGCACCCG	CAGATGACGC	ACCAGCAGAT	6150
	GACGCACCCG	CAGATGACGC	ACCAGCAGAT	GACGCAACAA	CATGTATCTC	6200
30	GAAAGGCTCT	TGTGGCTGGA	TGCGCTGCTC	GGATGACGAT	GACAAAAAAT	6250
	ACAAAAAAGC	ACTGAAAAAA	CTGGCAAAAC	TGCTGTGAATG	AGGGCGCCTG	6300
	GATCCCAATC	ACTTCTGGCT	AATAAAGATG	CAGAGCTCTA	GAGATCTGTG	6350
	TGTTGGTTTT	TTGTGGAATC	GCTGTGCCCT	CTAGTTGCCA	GCCATCTGTT	6400
	GTTTGCCCTT	CCCCGCTGCC	TTCTTGACC	CTGGAAGGTG	CCACTTCCAC	6450
35	TGTCCTTTCC	TAATAAAATG	AGGAAATGTC	ATCGCATTGT	CTGAGTAGGT	6500
	GCTCATCTAT	TCTGGGGGGT	GGGGTGGGGC	AGCACAGCAA	GGGGGAGAGT	6550
	TGGGAAGACA	ATAGCAGGCA	TGCTGGGGAT	GCGGTGGGCT	CTATGGGTAC	6600
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCTCT	6650
	CTCAGGGGGG	GGCCCGGTAC	CCAATTTCGC	CTATAGTGAG	TCGTATTACG	6700
40	CGCGCTCACT	GGCCGCTGTT	TTACAACGTC	GTGACTGGGA	AAACCTCTGG	6750
	GTTTACCCAAC	TTAATCGCCT	TGCAAGCAT	CCCCCTTCG	CCAGCTGGCG	6800
	TAATAGCGAA	GAGGCCCGCA	CGATTCGCC	TTCCCAACAG	TTGCGCAGCT	6850
	TGAATTCGGA	ATGGAATTTG	TAAGCGTTAA	TATTTTGTGA	AAATTCGCTG	6900
	TAAATTTTTG	TTAAATCAGC	TCATTTTTTA	ACCAATAGGC	CGAAATCGGC	6950
45	AAAATCCCTT	ATAAATCAAA	AGAATAGACC	GAGATAGGGT	TGAGTGTGTT	7000
	TCCAGTTTGG	AACAGAGTCT	CACATATATA	GAACGTGGAC	TCCAACCTCA	7050
	AAGGGCGAAA	AACCGTCTAT	CAGGGCGATG	GCCCACTACT	CCGGGATCAT	7100
	ATGACAAGAT	GTGATCCAC	CTTAACTTAA	TGATTTTTTAC	CAAAATCATT	7150
	AGGGGATTCA	TCAGTGCTCA	GGGTCAACGA	GAATTAACAT	TCCGTCAAGA	7200
50	AAGCTTATGA	TGATGATGTG	CTTAAAAACT	TACTCAATGG	CTGGTTATGC	7250
	ATATCGCAAT	ACATGCGAAA	AACCTTAAAG	AGCTTGCCGA	TAAAAAAGGC	7300
	CAATTTATTG	CTATTTACCG	CGGCTTTTTA	TTGAGCTTGA	AAGATAAATA	7350
	AAATAGATAG	GTTTTATTG	AAGCTTAAAT	TTCTTTATCG	TAAAAAATGC	7400
	CTCTTTGGGT	TATCAAGAGG	GTCAATTATAT	TTCCGGGAAT	AACATCATTT	7450
55	GGTGACGAAA	TAACTAAGCA	CTTGCTCTCT	GTTTACTCCC	CTGAGCTTGA	7500
	GGGGTTAACA	TGAAGGTCAAT	CGATAGCAGG	ATATAATATC	AGTAAAAACG	7550
	TAAACCAATA	ATCCAAATCC	AGCCATCCCA	AATTTGATGT	GAATGATTAT	7600
	AAATACACAG	AAACAGTAAAT	GGGCCAATAA	CACGGTGTGC	ATTGGTAAGG	7650
	CTCACCAATA	ATCCCTGTAA	AGCACTTCGC	TGATGACTCT	TTGTTTGGAT	7700
60	AGACATCACT	CCCTGTAAATG	CAGGTAAAGC	GATCCCAACA	CCAGCCCAATA	7750
	AAATTAANAAC	AGGGAANAAC	AACCAACCTT	CAGATATAAA	CGCTAAAAAG	7800

	GCAAAATGCAC	TACTATCTGC	AATAAATCCG	AGCAGTACTG	CCGTTTTTTC	7850
	GCCCCATTTA	GTGGCTATTG	TTCTGCCAC	AAAGGCTTGG	AATACTGAGT	7900
	GTAAAGAGCC	AAGACCCGCT	AATGAAAGC	CAACCATCAT	GCTATTCCAT	7950
	CCAAAACGAT	TTTCGTAATA	TAGCACCCAC	ACCGTTGCGG	GAATTTGGCC	8000
5	TATCAATTGC	GCTGAAAAAT	AAATAATCAA	AAAAATGGCA	TCGTTTATAA	8050
	TAAAGTGATG	TATACCGAAT	TCAGCTTTTG	TTCCCTTTAG	TGAGGGTTAA	8100
	TTGCGCGCTT	GGCGTAATCA	TGGTCATAGC	TGTTTCCTGT	GTGAAATTTG	8150
	TATCCGCTCA	CAATTCCACA	CAACATACGA	GCCGGAAGCA	TAAAGTGTA	8200
	AGCCTGGGGT	GCCTAATGAG	TGAGCTAACT	CACATTAAAT	CGCTTCCGCT	8250
10	CAGTCGCCCG	TTTCCAGTCG	GGAAACCTGT	CGTGCCAGCT	GCATTAAATG	8300
	ATCCGCCAAC	GGCGCGGGAG	AGGCGGTTTG	CGTATTGGGC	GCTCTTCCGC	8350
	TTCTTCGCTC	ACTGACTCGC	TGCCTCCGCT	CGTTCCGCTG	CGCGGAGCGC	8400
	TATCAGCTCA	CTCAAGAGCG	GTAATACGGT	TATCCACAGA	ATCAGGGGAT	8450
	AACGCAGGAA	AGAACATGTG	AGCAAAAGGC	CAGCAAAAGG	CCAGGAAACG	8500
15	TAAAAAGGCC	CGGTTGCTGG	CGTTTTTCCA	TAGGCTCCGC	CCCCCTGAGC	8550
	AGCATTCACA	AAATCGACGC	TCAAGTCAGA	GGTGGCGAAA	CCCGACAGAA	8600
	CTATAAAGAT	ACCAGGCGTT	TCCCGCTGGA	AGCTCCCTCG	TGCGCTCTCC	8650
	TGTTCCGACC	CTGCGCGTTA	CGGATACCTT	CTCCGCTTTT	CTCCCTTCCG	8700
	GTAAGCTGGC	GCTTTCTCAT	AGCTCACGCT	GTAGGTATCT	CAGTTCCGGT	8750
20	TAGGTCGTCT	GCTCCAAGCT	GGGCTGTGTG	CACGAAACCC	CCGTTGACCC	8800
	CGACCGCTGC	GCTTTATCCG	GTAACATTCG	TCTTGAGTCC	AACCCGCTGA	8850
	GACACGACTT	ATCGCCACTG	GCAGCAGCCA	CTGGTAACAG	GATTAGCAGA	8900
	GCAGAGTATG	TAGGCGGTGC	TACAGAGTTC	TGGAAGTGGT	GGCCTAATCT	8950
	CGGCTACACT	AGAAGGACAG	TATTTGGTAT	CTGCGCTCTG	CTGAAGCCAG	9000
25	TTACCTTCGG	AAAAGAGATT	GGTAGCTCTT	GATCCGGCAA	ACAAACCAAC	9050
	GCTGTGAGCG	GTGGTTTTTT	TGTTTGCAG	CAGCAGATTA	CGCGCAGAAA	9100
	AAAAGGATCT	CAAGAAGATC	CTTTGATCTT	TTCTACGGGG	TCCTGACGCT	9150
	AGTGAACGCA	AAACTCACGT	TAAGGGATTG	TGCTCATGAG	ATTATCAAAA	9200
	AGGATCTTCA	CCTAGATCCT	TTTAAATTAA	AAATGAAGTT	TTAAATCAAT	9250
30	CTAAAGTATA	TATGAGTAAA	CTTGGCTCGA	CAGTTACCAA	TGCTTAATCA	9300
	GTGAGGACCC	TATCTCAGCG	ATCTGTCTAT	TTCTGTTATC	CATAGTTGCC	9350
	TGACTTCCCG	TCGTGTAGAT	AACTACGATA	CGGAGGGGCT	TACCATTGCG	9400
	CCCCAGTGCT	GCAATGATAC	CGCGAGACCC	ACGCTACACG	GCTCCAGATT	9450
	TATCAGCAAT	AAACCAGCCA	GCCGGAAGGG	CCGAGGCGAG	AAGTGGTCTT	9500
35	GCAACTTTAT	CCGCTTCCAT	CCAGTCTATT	AATTGTTGCC	GGGAGACTAG	9550
	AGTAAGTAGT	TGCGCAGTTA	ATAGTTTGGG	CAACGTTGTT	GCCATTGCTA	9600
	CAGGCATCGT	GGTGTACAGC	TCGTGCTTTG	GTATGGCTTC	ATTGAGCTCC	9650
	GGTTCCTCAAC	GATCAAGGCG	AGTTACATGA	TCCCCCATGT	TGTGCAAAAA	9700
	AGCGGTTAGC	TCCTTCGGTC	CTCCGATCGT	TGTCAGAAAT	AAGTTTGGCG	9750
40	CAGTGTTATC	ACTCATGGTT	ATGGCAGCAC	TGCATAAATC	TCTTACTGTC	9800
	ATGCCATCCG	TAAGATGCTT	TTCTGTGACT	GGTGAGTACT	CAACCAAGTC	9850
	ATTCTGAGAA	TAGTGTATGC	GGCGACCGAG	TGTCTCTTGC	CCGCGCTCAA	9900
	TACGGGATAA	TACCGCGCCA	CATAGCAGAA	CTTTAAAGAT	GCTCATCAAT	9950
	GGAAAGACGTT	CTTCGGGGCG	AAACTCTCA	AGGATCTTAC	CGCTGTGAG	10000
45	ATCCAGTTCTG	ATGTAACCCA	CTCGTCACCC	CAACTGATCT	TCAGCATCTT	10050
	TTACTTTTCC	CAGCGTTTCT	GGGTGAGCAA	AAACAGGAAG	GCAAAATGCC	10100
	CCAAAAAAGG	GAATAAGGGC	GACACGGAAA	TGTTGAATAC	TCATACTCTT	10150
	GCTTTTTCAA	TATTATTGAA	GCAITTTATCA	GGGTTATTGT	CTCATGAGCG	10200
	GATACATATT	TGAATGTATT	TAGAAAAATA	AACAAATAGG	GSTTCCGCGC	10250
50	ACATTTTCCC	GAAGAGTGCC	AC			10272

SEQ ID NO:47 p1nMCS (CMV-CHOVg-ent-ProInsulin-synPA)

	1	ctgaagcgcc	ctgttagogg	gcattaagcg	cgsggggtgt	gggtgttacg	cgagcggtga
55	61	cogotacact	tgccagcgcc	ctagcgcccg	ctcctttcgc	ttctttctct	tcctttctct
	121	coagtttgcg	cggtcatoga	ttggctattg	gcaattgcac	agttgtgtac	catatcataa
	181	tatgtacatt	tatatgggt	catgtccaac	attacgcgca	tggtgacatt	gattattgac
	241	tagttattaa	tagtaatcaa	ttaaggggtc	attagttoat	agcccatata	tggaattcgc
	301	ogttacaata	cttaoggttaa	atggcccgcc	tggtgtgacc	cccaacgacc	cccgccattt
60	361	gacgtcaata	atgaogtagt	tttcccatagt	aacgcacata	gggactttcc	attgaogtca
	421	atgggtggag	tattttacgt	aaactgocca	cttggcagta	catcaaggtg	atcatatgoc
	481	aagtagccgc	cctattgaag	tcaatgaagc	taaatggccc	gocctggcatt	atgcccagta

	541	catgacctta	tgggaacttc	ctactttggca	gtacatctac	gtattagtca	togctattac
	601	catggtgatg	gggttttggc	agtaacatcaa	tgggtgggtga	tagcgggttg	actcaocggg
	661	attccaggt	ctccacccca	ttgaagctcaa	tgggaagcttg	ttttggccac	aaaatacaag
5	721	ggactttcca	aaatgtcgta	acaactcgcg	cccatgtgaog	caaatggggg	gtatgggtgtg
	781	acgggtggag	gtctataata	gcagagctcg	ttttagtgaac	cgtagatagc	ctgtggagacg
	841	ccatccacgc	tggttttgac	lccatagaag	acacoggagac	cgatccagcc	tcocggggcg
	901	ggaacgggtc	atggaaacgc	ggatctccgc	tgccacagagt	gaagtaagta	cogctatgac
	961	actctatagg	caaccccttt	tggctcttat	gaatgtatag	ttgttttttg	cttggggcct
10	1021	atacacccc	gcttctctat	gtctatagg	tttagctata	tttagctata	gggtgggggt
	1081	attgacctat	attgacacct	ccctcatggg	tgacgatcaat	ttccattact	aatccataac
	1141	atgggtcttt	ggcaacaacta	ctctctattg	ctatagtatg	actactctgc	cttcagagac
	1201	tgaacgggac	tgtatttttt	tacagagattg	gggtccattt	attattttac	aattccacata
	1261	tacacacacg	cogtccccgc	tgcccgaggt	ttttattaaa	catagctggg	gatctccogc
15	1321	ogaattctgg	gtaogtgttc	cggaacatgg	ctctctccg	gtagccggcg	gctgtccca
	1381	tcggagccct	ggctccatgc	cgacggcgc	tcattgtcgc	tcggcagctc	gtctctccta
	1441	acagttggag	cagagacttag	gcacagacaa	atgccacaac	ccaccagtg	gcgcgacaa
	1501	ggctggggcg	tagggtatgt	gtctgaaat	gaagctggag	attgggctgc	cagcgtgtgac
	1561	gcagatggaa	gacttaaggg	agcgccagaa	gaagatgcag	gcagctgagt	tgtctgattc
20	1621	tgatacaagt	cagaggtaac	tcocgtttgc	gtgctgttaa	cgggtggagg	tgctgtatgc
	1681	tgagcagtac	tcgttgtctg	cgccgcggcc	accagacata	atagctgaca	gaactacaga
	1741	ctgtctcttt	cagctggctct	tttctgcagt	cacctgtgga	ccatgtgoga	actcgatgat
	1801	ttacaagact	ctctttacaa	attctgcgcc	gaattacact	taaaacgact	caacagcttta
	1861	actttggctt	gcacgcatt	acttgactgt	aaaactctca	ctcttaccga	acttggcgtt
25	1921	aacctgccaa	ccaaagcgag	aacaaaact	aaacatcaaa	gactgcagcc	atttgttagt
	1981	aatcgtccac	tcacaaaaga	gogactcgct	gtatacogtt	ggatctgat	ctttactctg
	2041	tggggcaata	cgatgcccat	tgtacttgtt	gaactgctct	atattctgta	gcacaaaaga
	2101	cttatggtat	tgogagcttc	agtgctgact	caagctgtgt	ctgttaactc	ttatggagaa
	2161	gogttccgcg	tttcagagca	atgttcaaa	aaagctctat	accaatttct	agccgcacct
30	2221	gcgcgatttc	tacogagtta	caccacacgc	ctcatgttca	gtgatgctgg	ctttaaagct
	2281	ccatgggtata	aatcogttga	gaagctgggt	tggtactggt	taagtgcagt	aagagaaaaa
	2341	gtcacaatg	cagacocagg	agcggaacac	tggaacacta	tcagcaactc	actgatgatg
	2401	tcctctatgc	actcaaaagc	ttttaggctat	aagagcgctga	ctaaaagcaa	tcacattctca
	2461	tgccaaattc	tattgtataa	attcogtctc	aaagcgcgaa	aaaactcagc	cttcgacagc
35	2521	actcatgttc	acaccccgct	acctaaaact	tactcagctg	cggaacaggg	gcatgttgct
	2581	ctagcaacta	acttactctg	tgaatttoga	acacccaaac	aaactgttaa	tatctatggt
	2641	aagcgatgac	agattgaaga	aaacctccga	gaactgaaaa	gtcctgccta	cgagataggg
	2701	ctacgccata	gcgcaacggg	cagctccagag	ogttttgata	tcactgtctg	atacgcocctg
	2761	atgcttcaac	taacatgttg	gcttgogggc	gttcatgtct	agaaacaagg	ttgggacaa
40	2821	cacttccagg	ctaacacagt	cagaaaatga	aaagctactct	caacagttgc	cttaggcatg
	2881	gaagttttgc	ggcattctgg	ctacacaata	acaaggggaag	acttactcgt	ggctgcacac
	2941	ctatagactc	aaaattttat	cacacatgggt	taagctttgg	ggaaatttatg	agggatgcgc
	3001	tctagagoga	tcggggatct	oggggaaaagc	gttggtgacc	aaaggtgcct	tttatcatca
	3061	cttttaaaaat	aaaaaacatc	tactcagctgc	ctgttataag	cagcaatgac	ttatgatgta
45	3121	tgccatcatc	acacacaaaa	ctgattttaac	aaatgggttg	ttgcctctag	aaagtatatt
	3181	tgaacattat	cttgattata	ttattgataa	taataaaaaa	cttatcccta	tcacagaagt
	3241	gactgcatac	attgtgttga	atgaaactga	aaaaaaatga	acttgaatc	attactggta
	3301	aggttaacgc	catgtcagc	aaatgtatcc	aaagagaacca	accttaaagc	ttctcgacgc
	3361	aattgttaatt	ctgttgacc	ctgagcaactg	atgaatcccc	taagttaatt	ggtaaaaaat
50	3421	attaagttaa	gggtgataca	caactctgtca	tatgtatcog	taagtgtgag	ttagctcaat
	3481	cattagccac	ccagggcttt	acacttttatg	cttcogggctc	gtatgtgttg	tggaaattgtg
	3541	aggggataac	aatttccaac	aaagaaacgc	tatgacatg	atacggccaa	gcgcgcatt
	3601	aaactcact	aaagccacga	aaagctggag	ctccacccgc	gtggggggcg	ctctagaact
	3661	agtggaatccc	cogggatcca	gattggctatg	tgggccattgc	atacgtttga	ttcatatcat
55	3721	aattatgtaca	tttatattgg	ctcaggtcca	acattaccgc	catgttgaca	ttgatatttg
	3781	actagttatt	aattgataat	aattaggggg	ctattagttc	atagccocata	ttatggagtc
	3841	cgogttacat	aacttaaggt	aaatggccgc	ctgggtgac	cgcccaacga	cccccgccca
	3901	tttagctcaa	taatgaagta	tggtccocata	gtaacgcocaa	tagggacctt	tagctagct
	3961	ctnaggggtg	agattttacg	gtaactggac	cactttggac	tacatcaagt	gtatcatatg
60	4021	ccaagtacgc	ccctatttga	cgtaactgac	ggtaaatggc	cgcctggcca	ttatgccag
	4081	tacatgacct	tattggacct	tcctacttgg	tcactacatc	acgtattagt	catcgtattt
	4141	acactggtga	tgccgttttg	gcaatcacatc	aatggggcgtg	gaagcgggtt	tgactcacgg
	4201	ggattttccaa	gtctccacgc	ctctgaagtc	aatggggagt	tggtttggca	ccaaaaatcaa
	4261	cggaacttcc	caaaattgct	taacaaactc	ccoccaattga	ogcaaatggg	oggttagcgt
	4321	gtcaggtggg	aggtctatat	aagcagctgc	cggtttagta	acgctcagat	cgctgggaga
65	4381	cgccactccac	gctgctttga	ccctcataga	agacacccgg	accgatccag	tacccggcgc
	4441	cggaacagg	gcatgggaac	cggaattccc	gtgcacaga	gtgacgtaga	gggttggggc
	4501	agaactctata	ggcacaacccc	tttggctctt	atgcatgcta	taactgtttt	gggttggggc
	4561	ctatacaccc	cogcttctct	atgatatgta	tgatgggata	gotttagccta	taggtgtggg

	4621	ttattgacca	ttattgacca	ctccctatt	ggtagacata	ctttccatta	ctaatacata
	4681	acatggtctct	ttggccacac	tatctctatt	gggtatatgc	caatactctg	tcoctcagag
	4741	ctcgacacgg	actctgtatt	tttacaggat	ggggctcagg	ctttattatta	caaatctca
5	4801	tatacaaca	cgccgtcccc	cgtagccgca	gtttttatata	aacatagcgt	gggatctcca
	4861	cggaactctc	gggtacgtgt	tcgggaactg	ggctctctctc	cgtagccggc	ggagcttcca
	4921	cattcgagcc	ctgtgtcccat	gctccagcgt	gctcatgtgtc	gctcggcagc	ttcctgtctc
	4981	taacagctgga	ggccagacct	aggcacagca	caatgccacc	caccaccagt	gtccgcagca
	5041	agcccgctggc	ggtaggggtat	gtgtctgaaa	atgagcgtgg	agattgggct	cgccagcgctg
	5101	acgcagatgag	aagacttaag	gcagcggcag	aagaagatgc	agcgactgca	gtgtgtgtat
10	5161	tctgataaga	gtcagagggta	actcccgctg	cggtgtcgttt	aacggtggag	ggcaggtgatg
	5221	tctgagcagt	actcgttgtc	gcgcgcgcgc	ccaccagaca	taactgctga	cagacttaaca
	5281	gaactgtctc	ttccattgggt	cttttctgca	gtccacgtcg	gatcaatggg	ctccactcgt
	5341	gcagcagca	tggaaatttg	ttttgatgta	ttcaaggagc	tcgaactcca	ccatgccaat
	5401	gcagaacatct	tctactgcgc	cattgcoact	atgtcagctc	tagcctaagg	ataactcggg
15	5461	gaaaaagaca	gcaccaggac	acaaataaat	aaggttgttc	gcttataaa	acttccagga
	5521	ttcgagagaca	gtattgaagc	tcagtgtggc	acactctgaa	acgttcaact	ttcaacttaga
	5581	gacactctca	accaaataac	caaaccaaat	gatgtttatt	cggtcagcgt	ggcagctagg
	5641	ctttatgctg	aagagagata	cccaactcgt	ccagaatact	tgacgtgtgt	gaaggaaactg
	5701	taacagatgag	gcttgtgaac	tatacaactt	caaacagctg	agatcaagc	tcagcagcagc
20	5761	atcaactctc	gggtgaaag	tcagacaact	ggaattatca	gaaattgtct	tcagcagcagc
	5821	ctccgtgatt	ctcaaacctg	aatggttctc	gttaattgca	ttgtctccaa	aggactctgg
	5881	gagaaagcat	ttagaggatga	agacacacaa	gcaatgctct	tcagagtga	tgagcaagaa
	5941	agaaaacatg	tcgcagatgat	gtaccagatt	gggtttatta	gaagtggcat	aattgctctc
25	6001	gagaaataga	agatcctgga	gcttccattt	gccagtgagg	caatgagcat	gttggtgctg
	6061	tctgtctagt	aggtctcagg	octtgagcag	cttgagagta	taatacaact	tgaaacaactg
	6121	actgaatgga	ccagttctaa	tgttatggaa	gagagaagat	caaaagtgtac	ttactctgca
	6181	tgaagatgga	ggaaaaatac	aaactccact	ctcgtctaat	ggctatgggc	attaactagc
	6241	ttgttagctgc	ttcagccaat	ctctctggca	tctcctcagc	agagagccctg	agaatatctc
30	6301	aagttctgca	tgagcacact	gcagaaaata	atgaagcagg	cagagaggtg	gttagggtagc
	6361	acagcctggt	agtggtgatg	gcagagctct	ctgaagaatt	taagggtgac	catcaactcc
	6421	tctctgttat	caagacacat	gcgaacaaag	ccgtttctctc	cttttggaag	atgtgtttcc
	6481	cgccggccagc	agatgaagca	ccagcagatg	acgcacacgc	agatgaagca	ccagcagatg
	6541	acgcacacgc	agatgaagca	acaaacatga	tcctgaagag	octttgtggc	tgagctggcc
35	6601	tgcttgagta	cgatgacaaa	tttctgaaac	aaacactgtg	cggtctcac	ctgggtgaag
	6661	ctctctcaact	agtgctgggg	acacagaggt	tctctcaaac	acccaaagac	cgccggggag
	6721	cagaggacact	gcaggtgggg	ccctggagga	tgggcggggg	ccctgggtgc	ggacagctgc
	6781	gcactgtgct	octggagggg	ccctggagga	agcgtggcat	tgtagaatac	tgctgttaaca
	6841	atcgccggcg	ctctagacaa	ggcgcttgga	tcagatcaac	ggcgccctaa	agggcggaatt
40	6901	gagctctaga	gatctgtgtg	ttggtttttt	ttggaatctg	ttgtcctctc	agttgcagcc
	7021	catctgttgt	ttgcccctcc	cccgctcctt	ccctgaacct	ggaaggtgcc	actcccaact
	7081	tcctttccta	ataaaatgag	gaaattgcac	gcattgtgtc	gagtaggtgt	catctcttc
	7141	tggggggtgg	ggtagggcag	caacgcaagg	ggggagtagt	ggaagacaat	agcaggaact
	7201	ctggggatgc	ggtagggctc	atgggtacct	ctctctctct	ctctctctct	ctcagctctc
45	7261	ctctctctcg	gtacctctcc	tcgagggggg	gcocgttaac	caatttgcgc	tatatgtagt
	7321	gctattacgc	gcgcctcaatg	gcgcgtgttt	tacaaactgc	tgactgggaa	aaacctgggt
	7381	ttccccaact	taatgccttc	gcagcacact	ccctcttgcg	cagctggggt	aatagogaag
	7441	agcccgccac	cgatgccttc	ttcccaacag	tgccagcgtc	gaatggcgaa	tggaatttgt
50	7501	aagcgttaat	attttgttaa	aatttgcgtt	aaatttttgt	taaatcagct	acttttttaa
	7561	ccaatagcc	gaaatggcca	aaatccctta	taaatcaaaa	gaatagacg	agatagggat
	7621	gaggtgttgt	ccagtttgga	acagagctcc	actataaag	aacgtggact	ccaaactgaa
	7681	agggcgaaaa	acogtctcat	aggcgatagg	cccaactact	cgggatcata	tgacaagatg
	7741	tgatccacc	ttacttaact	gatttttaac	aaaactcata	ggggatcatc	cagtgctcag
55	7801	gctcaacag	aattaacatt	ccgtcaggaa	acttatgatc	gatgatgtgc	ttaaaaactt
	7861	actcaatggc	tggttatgca	tatogcaata	catcgogaaa	acctaamaag	gcttgccgat
	7921	aaaaaagcc	aatttattgc	tatttaacgc	ggctttttat	tgagcttgaa	agatacaata
	7981	aatagtagg	ttttatttga	agctaaatct	tottttatgt	aaaaaatgc	ctcttggtgt
	8041	ctcaagagg	tcattatatt	tcgcggaaat	acatcatttg	gtgacgaat	aactcaagac
60	8101	ttgtctcctg	tttaactccc	tgagcttgag	gggttaacat	gaaggtcatc	gatagcagga
	8161	taataataca	gtaaaaacgt	aaaccaaata	tcocaaatca	gocactccaa	atttggtagt
	8221	aatgatata	ataaacagca	aacagtaagt	ggcocaatac	acoggttgca	ttcgtagaagc
	8281	tcocaaata	tcocgtgaaa	gcacottgtc	gatgactctt	tgtttggata	gacatacact
	8341	ctctgtaatg	aggttaaaog	atcccaacc	cagcocaata	aattaaaaac	gggaataact
	8401	accacacctc	agatataaac	gctaaaaaag	actatgcaat	actatctgca	ataactcga
65	8461	gcagtagctg	ogtttttttg	ccattattagt	ggctattctt	octcgcacaa	agggctggaa
	8521	taactagtggt	aaaagaccaa	gaccogtaac	aaagagocaa	ccatcatgct	attcactcat
	8581	acgattttctg	taatagcaac	acacogtctg	ggatttgcta	tcactgogct	gaataataaa
	8641	gcacaaatg	gcacogttaa	ataagtgatg	tatacogcatc	agcttttgtt	gaatttagtg

	8701	agggttaatt	gcgcgcttgg	cgtaatcatg	gtcatagctg	tttcctgtgt	gaattgttta
	8761	tcgcgtcaac	atctccacac	acatacagac	cggaagataa	aagtgtaaag	ccctgggtgco
	8821	ctaagtgtag	agctaacctca	cattaattgc	gttgcgtcca	ctgcgcctgc	tcocagtggcg
	8881	aaactgtgtg	tgccagctgc	attaatgaat	cgcccaacgc	gcggggagag	gcggtttgctg
5	8941	tattggsgcg	tcttcogctt	cctgcgtcac	tgactcgtct	cgctcggtgc	ttcgggtgcg
	9001	gcagcgctca	tcagctcaact	caaaggcggt	aatacgttta	tcacagaaat	cagggtataa
	9061	cgagggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	aggaacgcga	aaaaggccgc
	9121	gtgtcgtggc	tttttccata	ggctcgcgcc	ccctgacgag	catcacaaaa	atcgacgctc
10	9181	aagtccagag	tgcgaaaccc	cgacaggact	ataagatac	caggcgtttc	ccctgggaag
	9241	ctccctcgtg	cgctctcctg	ttccgacctc	gcgcgttacc	ggatacctgt	ccgctttctt
	9301	cccttcggga	agcgtggcgc	ttctctatag	ctcacgtctg	aggtatctca	gttcgggtga
	9361	ggctcgtcgc	tccaagctgg	gctgtgtgca	cgaaaccccc	gttccgccc	accgctggcg
	9421	cttatccoggt	aactatcgct	ttgagtccaa	cccggtaaga	cacgaactat	cgccaactggc
	9481	agacgcacat	ggtaacagga	ttagcagagc	gaggtatgta	ggcgggtgcta	cagagtctct
15	9541	gaagtcgtgg	cctaactacg	gctacactag	aaggacagta	tttggtatct	gcgctctgct
	9601	gaagcagact	accttcggaa	aaagagttgg	tagctcttga	tcgggcaaac	aaacacccgc
	9661	tggtagcggg	ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca
	9721	agaagatcct	ttgatctttt	ctacggggct	tgacgtctag	tggaacgaaa	actcaogtta
	9781	agaggttttg	gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	taaatataaa
20	9841	cttaagctttt	aaatcaatct	aaagtataata	tgagtaaaat	tggtctgaca	gttacaagtg
	9901	cttaactcagt	gaggcaacct	tctcagcgat	ctgtctatct	cgttcatoca	ccagtctgct
	9961	actcccccgc	gtgtagataa	ctcacgatac	ggaggggtta	ccatctggcc	ccagtctgct
	10021	aattgataccg	cgagaccacac	gctcacccgc	tcacagattta	tcagcaataa	accagccagc
25	10081	cggaaggccc	gagcgcagaa	gtggtcctgc	aaacttatcc	gcctccatcc	agcttatata
	10141	ttgttgccgg	gaagctagag	taagtagttc	gcacgttaat	agtttgccga	acgttggtgc
	10201	catctgtaca	ggcatcgtgg	tgctacgctc	gtcgtttggt	atggtctcat	tcagctccgg
	10261	ttcccaacga	tcaggcgagc	ttacatgata	ccccatggtt	tgcaaaaaag	cggttagctc
	10321	cttcggtcct	cagatcgttg	tcagaagtaa	gttgcccgca	gtgttatcac	tcagtgggtat
	10381	ggcagcactg	cataaattctc	ttactgtcat	gcacatcgta	agatgctttt	ctgtgactgg
30	10441	tgagtaactca	accaagtcac	tctgagaata	gtgtatcggg	cgaccagagt	gctcttgccc
	10501	ggcgtcaata	cgggataata	ccgcgccaca	tagcagaact	ttaaaagtcg	tcacatctgg
	10561	aaaaogttct	tcggggcgaa	aactctcaag	gacattacgc	ctgttgagat	ccagttcgat
	10621	gtaacccact	cgtgcaccca	actgatcttc	agcatctttt	actttccaca	gcgtttctgg
	10681	gtgagcaaaa	acagggaagg	aaaatgcgcg	aaaaaaggga	ataaggcgca	caaggaaatg
35	10741	ttgaatactc	atactcttcc	tttttcaata	ttatgtgaag	attattcagg	gttatgtctc
	10801	catgagcgga	tacataattg	aatgtattta	gaaaaataaa	caaatagggg	ttccgscac
	10861	atttcccgga	aaagtgcacc				
40	SEQ ID NO:48 (cecropin prepro)						
	AAT TTC TCA AGG ATA TTT						
	TTC TTC GTG TTC GCT TTG						
	GTT CTG GCT TTG TCA ACA						
	GTT TCG GCT GCG CCA GAG						
45	CCG AAA						
	SEQ ID NO:49 (cecropin prepro extended)						
	AAT TTC TCA AGG ATA TTT						
50	TTC TTC GTG TTC GCT TTG						
	GTT CTG GCT TTG TCA ACA						
	GTT TCG GCT GCG CCA GAG						
	CCG AAA TGG AAA GTC TTC						
55	AAG						
	SEQ ID NO:50 (cecropin pro)						
	GCG CCA GAG CCG AAA						
60							
	SEQ ID NO:51 (cecropin pro extended)						
	GCG CCA GAG CCG AAA TGG AAA GTC TTC AAG						
65							
	SEQ ID NO:52 (a Kozak sequence)						
	ACCATGT						

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/41335

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12N 15/00, 15/63; C12P 21/00; A01K 67/027

US CL : 435/320.1, 455; 800/4, 7, 13, 14, 19, 21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 455; 800/4, 7, 13, 14, 19, 21

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WEST, STN, MEDLINE, CAPLUS BIOSIS, SCISEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HORN et al. A Versatile Vector Set for Animal Transgenesis. Development Genes and Evolution. 2000, Vol. 210, No. 12, pages 630-637, especially abstract, introduction.	1-20, 23-28
A	EGGLESTON et al. A sensitive and Rapid Assay for Homologous Recombination in Mosquito Cells: Impact of Vector Topology and Implications for Gene Targeting. BMC Genetics. 17 December 2001, Vol. 2, No. 21, http://www.biomedcentral.com/1471-2156 , online, pages 1-9, especially background, materials and methods.	1-20, 23-28
A	WO 01/71019 A1 (SAVAKIS et al.) 27 September 2001 (27.09.2001), especially, pages 10, 11, 27, 31.	1-20, 23-28

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "B" earlier application or patent published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search

07 October 2004 (07.10.2004)

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703) 305-3230

Date of mailing of the international search report

09 NOV 2004

Authorized officer

Shu-Lin Chen
Telephone No. 371-272-1600

INTERNATIONAL SEARCH REPORT

PCT/US03/41335

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claim Nos.: 21, 22
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐

☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.